

*Aus dem Institut für Chirurgische Forschung  
des Fachbereichs Medizin der Philipps-Universität Marburg*



## **Diagnostik des hepatozellulären Karzinoms – Eine Metaanalyse**

Inaugural-Dissertation zur Erlangung des Doktorgrades der gesamten  
Humanmedizin

dem Fachbereich Medizin der Philipps-Universität Marburg

vorgelegt von

Janina Pado aus Stuttgart

Marburg, 2015

Angenommen vom Fachbereich Medizin der Philipps-Universität Marburg  
am: 30.11.2015

Gedruckt mit Genehmigung des Fachbereichs.

Dekan: Professor Dr. H. Schäfer

Referent: Priv.-Doz. Dr. H. Sitter

1. Korreferent: Priv.-Doz. Dr. J. Tebbe

Originaldokument gespeichert auf dem Publikationsserver der  
Philipps-Universität Marburg  
<http://archiv.ub.uni-marburg.de>



Dieses Werk bzw. Inhalt steht unter einer  
Creative Commons  
Namensnennung  
Keine kommerzielle Nutzung  
Weitergabe unter gleichen Bedingungen  
3.0 Deutschland Lizenz.

Die vollständige Lizenz finden Sie unter:  
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

# Inhaltsverzeichnis

	Seite
1. Einleitung.....	1
1.1 Epidemiologie .....	1
1.2 Ätiologie.....	3
1.3 Pathologie.....	5
1.4 Symptomatik.....	6
1.5 Prävention .....	8
1.5.1 Primärprävention .....	8
1.5.2 Sekundär- und Tertiärprävention.....	8
1.6 Diagnostik .....	9
1.6.1 Index-Test und Referenzstandard .....	9
1.6.2 Ultraschall (US) .....	9
1.6.3 Computertomographie (CT) .....	10
1.6.4 Magnetresonanztomographie (MRT).....	11
1.6.5 Alpha-Fetoprotein (AFP).....	12
1.6.6 Kontrastmittel.....	12
1.6.7 Feinnadelbiopsie.....	14
1.7 Staging .....	15
1.7.1 TNM-Staging-System .....	15
1.7.2 Okuda-Staging-System .....	16
1.7.3 Cancer of the Liver Italian Program (CLIP) -Staging- System.....	16
1.7.4 Japanese Integrated Staging Score (JIS) .....	17
1.7.5 Barcelona Clinic for Liver Cancer-Staging-System (BCLC) .....	17
1.8 Therapie.....	18
1.8.1 Tumorresektion .....	19

1.8.2 Lebertransplantation .....	19
1.8.3 Radiofrequenzablation (RFA) .....	21
1.8.4 Transarterielle Chemoembolisation (TACE) .....	21
1.8.5 Systemische Therapie .....	22
1.8.6 Symptomatische Therapie .....	22
2. Fragestellung und Zielsetzung dieser Arbeit .....	24
3. Material und Methoden .....	25
3.1 Basisliteratursuche .....	25
3.2 Diagnostische Studien und Ultraschall .....	26
3.3 Diagnostische Studien und CT .....	27
3.4 Diagnostische Studien und MRT .....	27
3.5 Kombinierte Suche diagnostischer Studien .....	27
3.6 Diagnostische Studien und AFP .....	27
3.7 Zuteilung des Evidenzgrades .....	28
3.8 Bias .....	29
3.8.1 Spectrum Bias .....	29
3.8.2 Selection Bias .....	30
3.8.3 Verification Bias .....	30
3.8.4 Diagnostic Review Bias .....	30
3.9 Vorbereitung für die Analyse .....	31
3.10 Bivariate logistische Regression .....	33
3.10.1 Formeln und Vorgehen .....	34
3.11 Patient-by-patient- und Lesion-by-lesion-Analyse .....	36
4. Ergebnisse .....	37
4.1 Basisliteratursuche .....	37
4.2 Diagnostische Studien und Ultraschall .....	38
4.3 Diagnostische Studien und CT .....	39

4.4 Diagnostische Studien und MRT .....	40
4.5 Kombinierte Suche diagnostischer Studien .....	41
4.6 Auseinandersetzung mit den gefundenen Studien .....	41
4.7 Eigenschaften der aufgenommenen Studien .....	42
4.7.1 Studien mit Ultraschall .....	42
4.7.2 Studien mit Computertomographie .....	43
4.7.3 Studien mit Magnetresonanztomographie .....	45
4.7.4 Studien mit AFP .....	46
4.8 Daten der aufgenommenen Studien .....	48
4.8.1 Tabelle 2: Daten der Studien mit US als diagnostischem Test .....	48
4.8.2 Tabelle 3: Daten der Studien mit CT als diagnostischem Test .....	52
4.8.3 Tabelle 4: Daten der Studien mit MRT als diagnostischem Test .....	56
4.8.4 Tabelle 5: Daten der Studien mit AFP als diagnostischem Test .....	61
4.9 Ergebnisse der Meta-Analyse .....	63
4.9.1 Ergebnisse für die Diagnostik mit US .....	63
4.9.2 Ergebnisse für die Diagnostik mit CT .....	65
4.9.3 Ergebnisse für die Diagnostik mit MRT .....	67
4.9.4 Ergebnisse für die Diagnostik mit AFP .....	69
4.10 Impaktfaktoren .....	69
5. Diskussion .....	72
5.1 Interpretation der Ergebnisse und Vergleich der Subgruppen-Analyse .....	72
5.2 Fazit .....	75
5.3 Vergleich der Ergebnisse mit der Analyse von <i>Colli et al. (2006)</i> .....	76
5.4 Vergleich der diagnostischen Richtigkeit mit Ergebnissen anderer Autoren .....	77
5.5 Potentielle Limitationen dieser Meta-Analyse .....	83
6. Zusammenfassung .....	86
7. Abstract .....	88

8. Anhang.....	90
8.1 Anlage 1: Tabelle 16 - Impaktfaktoren und Evidenzlevel der Studien .....	90
8.2 Anlage 2: Fragebogen zur Ermittlung des Evidenzgrades .....	92
8.3 Anlage 3: Tabelle 17 – Ausschlussgründe für gefundene Studien .....	93
8.4 Anlage 4: Tabelle 18 – Levels of Evidence [60].....	98
9. Literatur .....	101
9.2 In der Metaanalyse verwendete Studien .....	107
10. Abkürzungsverzeichnis .....	113
11. Abbildungsverzeichnis .....	115
12. Tabellenverzeichnis .....	116
13. Verzeichnis der akademischen Lehrer.....	117
14. Danksagung .....	118

# 1. EINLEITUNG

## **1.1 Epidemiologie**

Jährlich erkranken mehr als 12 Millionen Menschen weltweit an Krebs [73], davon sind etwa 700.000 Fälle Leberkrebs. In mehr als 500.000 dieser Fälle handelt es sich um das hepatozelluläre Karzinom (HCC) [25], das somit die fünfhäufigste Krebserkrankung bei Männern und die achthäufigste bei Frauen darstellt [25], [83]. Bezogen auf Krebserkrankungen stellt das HCC die dritthäufigste Todesursache weltweit dar [83].

Hierbei finden sich Unterschiede zwischen einzelnen Staaten. In der westlichen Welt rangiert das HCC auf Platz 10 der häufigsten Malignome der Männer, während es bei den Frauen nicht einmal mehr unter den 10 häufigsten Krebsarten zu finden ist. Hingegen ist das hepatozelluläre Karzinom in den Entwicklungsländern der dritthäufigste Krebs bei Männern und immerhin der sechsthäufigste bei Frauen [40].

Mit etwa Dreiviertel aller Fälle weltweit am stärksten betroffen sind die Länder im Osten und Südosten Asiens [12], [13], wie Thailand, Japan oder Korea. Über die Hälfte aller HCC-Patienten weltweit leben in China [73] mit einer Inzidenz von bis zu 80 pro 100.000 [83].

An zweiter Stelle der Regionen mit der höchsten Inzidenz an HCC steht Afrika mit 21 Fällen pro 100.000 Einwohner bei den Männern und ca. 11 pro 100.000 bei den Frauen [83]. Insbesondere Staaten wie Tschad oder Kongo in Mittelfrika oder Mali im Westen des Kontinents fallen durch hohe Fallzahlen auf [12], [73]. Hingegen sind der Norden Europas, Neuseeland und Australien, sowie die kaukasische Bevölkerung Nordamerikas am wenigsten von Leberkrebs betroffen. In Nordeuropa liegt die Inzidenz bei Männern bei 3 pro 100.000 und bei Frauen bei 1 pro 100.000 [83].

Nahezu überall auf der Welt sind Männer häufiger von Leberkrebs betroffen als Frauen [13]. In Europa finden sich die größten Unterschiede mit einem Verhältnis von Mann zu Frau bis zu 8,8:1 [55]. In Regionen, die insgesamt eine hohe Inzidenz an HCC aufweisen, ist auch das Geschlechterverhältnis nicht ganz so ausgeprägt. So ist in Asien im Schnitt mit einem Verhältnis von 3:1 zu rechnen. Wohingegen in Südamerika



annähernd gleiche Verhältnisse zwischen Männern und Frauen zu finden sind [55]. Die Ursache für diese unterschiedliche Verteilung unter den beiden Geschlechtern ist wahrscheinlich in dem bei Männern häufiger vorhandenen risikobehafteten Lebensstil zu suchen. So setzen sich Männer öfter als Frauen den Risikofaktoren für die Entstehung von Leberzirrhose aus, wie dem Genuss von Alkohol oder intravenösem Drogenkonsum (und dem damit einhergehendes Risiko für die Infektion mit Hepatitis B- oder Hepatitis C-Viren) [27], [55].

Das HCC als histologisch wichtigster Subtyp ist mit etwa 75% aller Leberkrebsarten die häufigste primäre maligne Lebererkrankung [40]. Es zeichnet sich vor allem durch eine hohe Letalität aus. Die 5-Jahres-Überlebensrate von HCC-Patienten beträgt 6,5 % [25]. Die meisten Patienten sterben innerhalb eines Jahres nach Diagnosestellung.

Seit einigen Jahren lässt sich in der Inzidenz des hepatozellulären Karzinoms eine Zunahme erkennen [14], [25], so zeigt z.B. Westschottland eine Zunahme der HCC-Erkrankungen von ca. 4 pro 100.000 Männern im Jahr 1975 auf ca. 16 pro 100.000 Männern im Jahr 1990 [26]. Für Japan wird eine Zunahme der HCC-Inzidenz von 15-20% zwischen den frühen 1970er Jahren und 1999 angegeben [12]. Auch für die nächsten Jahre wird mit steigenden Zahlen von HCC-Fällen gerechnet [80]. Erklären lässt sich dies einerseits durch die zunehmende Anzahl von Zirrhose-Patienten sowie von mit dem Hepatitis-B- bzw. Hepatitis-C-Virus Infizierten [40], andererseits durch die zunehmend besseren Behandlungsmöglichkeiten mit damit einhergehender Lebensverlängerung für die Patienten [25].

Lediglich in sehr stark betroffenen Regionen Chinas wurde ein Rückgang der Inzidenz des HCC festgestellt; Erklärungen hierfür sind allerdings nicht eindeutig. Diskutiert werden der Rückgang von Risikofaktoren wie Hepatitis-Infektionen und Aflatoxin-Belastung in Lebensmitteln, sowie Änderungen in der Therapie und Prognose der Leberzirrhose oder Veränderungen in der Diagnostik und dem Screening bei HCC-Erkrankung [54].

## **1.2 Ätiologie**

Das hepatozelluläre Karzinom entsteht in 80% - 90% der Fälle im Rahmen einer Leberzirrhose [28]. Diese ist, unabhängig von ihrer eigenen Entstehungsursache, einer der Hauptrisikofaktoren für die Entwicklung von Leberkrebs. Die weltweit häufigste Ursache für eine Leberzirrhose ist die chronische Infektion mit Hepatitis B oder C. Etwa 350 Millionen Menschen auf der ganzen Welt sind chronisch mit dem Hepatitis-B-Virus (HBV) infiziert und immerhin ca. 170 Millionen mit dem Hepatitis-C-Virus (HCV) [55], [73].

Auch ohne das Vorliegen einer Zirrhose ist der HBV eine der wichtigsten Ursachen für die Entstehung von HCC [27]. Das Hepatitis-B-Virus wird auf dem Blutweg übertragen, weltweit am häufigsten infizieren sich Neugeborene bei der Geburt durch den Kontakt mit Blut der infizierten Mutter. Seltener ist der Infektionsweg im Erwachsenenalter über ungeschützten Sexualkontakt oder intravenösen Drogenkonsum. Zwar gibt es eine Impfung gegen HBV, jedoch wird diese vor allem in den Teilen der Welt, in denen diese Infektion weit verbreitet ist, nicht flächendeckend durchgeführt, sodass kein ausreichender Schutz der Bevölkerung besteht und somit immer noch große Endemiegebiete existieren.

Die Infektion mit HCV findet ebenfalls parenteral statt, d.h. durch den Kontakt mit infizierten Körperflüssigkeiten. Die Erkrankung heilt in ca. 50% der bei der Geburt infizierten Kinder ohne spezifische Behandlung spontan aus [73]. Entscheidend für das Fortschreiten zur Zirrhose scheinen hierbei nicht die Viruslast oder sonstige virale Parameter zu sein, sondern andere Faktoren, wie z.B. Alter, Geschlecht oder Alkoholkonsum [27]. Im Gegensatz zu HBV gibt es keinen Impfstoff gegen das Hepatitis-C-Virus.

Ein anderer Risikofaktor für die Entstehung von HCC ist das Gift des Schimmelpilzes *Aspergillus*. Mit Aflatoxin B1 infizierte Patienten haben ein vierfach erhöhtes Risiko, an einem hepatozellulären Karzinom zu erkranken, Kombinationen aus Aflatoxin B1 und HBV erhöhen dieses Risiko noch auf das Sechzigfache [55]. Dieses Gift wird mit dem Verzehr von mit Schimmelpilz befallenen Lebensmitteln im Körper aufgenommen. Es wird diskutiert, ob andere leberschädigende Toxine, wie zum Beispiel Vinylchlorid, einen relevanten Risikofaktor für die HCC-Entstehung darstellen – die geringe Fallzahl

und die unterschiedlichen Tumorarten, die hier entstehen (z.B. Gallengangstumoren, Gallenblasentumoren oder HCC) lassen diesen Zusammenhang allerdings nur schwer bestätigen [27].

Auch der Konsum von Alkohol gilt als begünstigend für die Entstehung eines Leberkrebses. Einerseits verstärkt Alkohol das Auftreten von HCC bei Patienten, die bereits eine HCV- oder HBV-Infektion aufweisen, und erhöht somit das bereits vorhandene Risiko dieser Patienten. Andererseits begünstigt der regelmäßige und starke Alkoholkonsum die Entstehung von Leberzirrhosen und insofern indirekt auch die HCC-Entstehung. Eine auf den Alkohol allein zurückgeführte Entstehung von Leberkrebs ohne das gleichzeitige Vorliegen einer Zirrhose oder Hepatitis-Infektion wird diskutiert [55]. Eine große Bedeutung bei der Entstehung von HCC kommt dem Alkohol vor allem in Europa und den USA zu, wohingegen er in Asien und Afrika in dieser Hinsicht eine eher untergeordnete Rolle spielt [83].

Als Risikopatienten gelten jene, deren Grunderkrankung das Auftreten von HCC wahrscheinlicher werden lässt. Dazu zählen alle Zirrhose-Patienten unabhängig von der Entstehungsursache, sowie alle Patienten mit chronisch rezidivierender Hepatitis-B-Infektion. Das Risiko innerhalb eines Jahres ein HCC zu entwickeln liegt bei Zirrhotikern zwischen 2,5%-3% [14], [55]; bei Hepatitis-B-infizierten Männern liegt das Lebenszeitrisiko einer HCC-Erkrankung zwischen 10%-25% [55]. Die Unterschiede in der Ätiologie und damit einhergehend in der Verteilung der HCC-Erkrankung weltweit führen zu einem unterschiedlichen Vorgehen in Diagnostik und Therapie zwischen westlichen und östlichen Ländern. So werden z.B. in östlichen Staaten häufiger kleinere Leberresektionen durchgeführt als im Westen [74].

Abgesehen von der alkoholinduzierten Leberzirrhose gibt es auch noch die sogenannte nicht-alkoholische Steatohepatitis (NASH). Hierbei kommt die Zirrhose durch eine Fettspeicherstörung in den Leberzellen und eine Entzündungsreaktion über die Entstehung einer Fibrose zustande [78]. Dies geschieht meist im Rahmen des metabolischen Syndroms. In Kombination mit Übergewicht ist die NASH zusammen mit der alkoholinduzierten Fettleber wahrscheinlich die Hauptursache für die Entstehung von Leberkrebs in der westlichen Welt [40]. Außerdem ist ein Zusammenhang zwischen Insulinresistenz bei Typ-I-Diabetikern und der Entstehung

von Zirrhose gesichert [27], [55], sodass auch der Glukosestoffwechsel eines Typ-II-Diabetikers im Verdacht steht, eine Zirrhoseentstehung zu begünstigen [13], [27], [55].

Im Vergleich sind die Hämochromatose, der  $\alpha$ 1-Antitrypsinmangel oder das autoimmune Geschehen als gesicherte Ursache für eine Zirrhoseentstehung von untergeordneter Bedeutung [27]. Ebenso wird erforscht, inwiefern Übergewicht und die daraus resultierende Veränderung in der Besiedlung mit Darmkeimen zu einer HCC-Entstehung führen kann [82]. Diskutiert werden unter anderem außerdem die langjährige Einnahme von oralen Kontrazeptiva oder das Rauchen von Zigaretten [27], [55] als Risikofaktoren für die Entstehung und das Trinken von Kaffee [27] oder grünem Tee [55] als Schutzfaktoren vor der Entstehung von HCC.

### **1.3 Pathologie**

Das hepatozelluläre Karzinom entsteht am Anfang meist als einzelner Knoten in der Leber. Dies kann prinzipiell überall in der Leber der Fall sein, überwiegend jedoch finden sich die Krebsgeschwüre im rechten Leberteil. Von diesem Zeitpunkt an können entweder neue Knoten hinzukommen oder das HCC streut über das venöse System, im Regelfall über die Portalvene [55]. Auch bei dem sogenannten diffusen HCC wird davon ausgegangen, dass es sich anfangs um nur einen Knoten gehandelt hat [55]. Charakteristisch für das HCC ist seine Durchblutung, welche fast ausschließlich arteriell und so gut wie nie über Portalvenenäste gewährleistet ist [31]. Dies hat unterschiedliche Folgen sowohl für Diagnostik als auch Therapie des HCC, wobei man versucht sich diese Tatsache zunutze zu machen.

Die genetische Vielfalt des HCC ist außerordentlich groß, sodass eine genaue Untersuchung im Hinblick auf Rezidivwahrscheinlichkeit und Ansprechen auf Therapie sinnvoll erscheint [22]. Molekulargenetisch lassen sich in den Leberzelltumoren viele unterschiedliche Mutationen nachweisen. Es sind verschiedene Gene an unterschiedlichen Stellen betroffen, sodass Veränderungen in den Reparaturmechanismen der DNA, sowie Störungen der Signaltransduktionswege und der Apoptose- und Wachstumsmechanismen der Zelle zu finden sind [69]. Im Allgemeinen kann man die für die Entstehung von Tumoren verantwortlichen Gene in zwei Gruppen unterteilen: Die erste Gruppe bilden jene Gene, deren Signalwege in der

gesunden Zelle zur Unterdrückung von tumortypischer Veränderung führen. Die zweite Gruppe besteht aus jenen Genen, die in der Zelle die Signalwege beeinflussen, welche zu Zellwachstum, Zellteilung und anderen Vorgängen führen, die bei Tumorzellen mit außergewöhnlichem Wachstum und schneller Verbreitung einhergehen. Die erste Gruppe wird auch als *Tumorsuppressorgene* bezeichnet; die zweite Gruppe als *Protoonkogene*. Werden die Allele, auf denen diese Gene liegen, durch Mutation verändert (bei Tumorsuppressorgen v.a. durch Eliminierungen und damit einhergehendem Funktionsverlust, im Fall von Protoonkogenen v.a. durch Mutationen, z.B. Punktmutationen [72]), so können die betroffenen Zellen sich zu Tumorzellen entwickeln. Speziell auf die Leber bezogen sind bereits folgende Gene bekannt, die für die Entstehung von Tumoren kritisch sind: p16-INK4A/p14-ARF, FHIT, AXIN1, p53 Gene, PTEN, NF2, STK11, BAX, LRPDIT [61], sowie  $\beta$ -Catenin und APC [69]. Laut *Spangenberg et al. (2004)* und *McGlynn et al. (2005)* sind außerdem Enzyme wie z.B. die Glutathion-S-Transferasen zu erwähnen, da sie karzinogene Stoffe wie das Aflatoxin B1 verstoffwechseln und somit ein Risiko für die HCC-Entstehung bergen können[55], [69].

Diese große genetische Variabilität erschwert die Möglichkeit eines Früherkennungstests oder eines vereinheitlichten molekulargenetischen HCC-Diagnostetests [69].

## **1.4 Symptomatik**

Die Symptome von Patienten mit HCC können sehr unterschiedlich sein. Da diese Erkrankung in den meisten Fällen mit einer ausgeprägten Leberzirrhose einhergeht, sind vor allem die dadurch bedingten Veränderungen augenscheinlich. Typische Zeichen der Leberzirrhose sind Schmerzen im rechten Oberbauch, Aszites, Palmarerythem, Caput medusae, Spidernaevi, Lackzunge und Ödeme [8], [25]. Hinzu kommen allgemeine Symptome wie ein schweres Krankheitsgefühl, Erschöpfung und Leistungseinbuße. Zudem kann durch die gestörte Entgiftungsfunktion der Leber eine hepatische Enzephalopathie eintreten [9]. Außerdem betroffen ist auch die Synthesefunktion des Organs, wodurch bestimmte Gerinnungsfaktoren vermindert produziert werden und es so zu einer vermehrten Blutungsneigung kommen kann. Allen Symptomen ist gemein, dass sie erst im späten Stadium der Zirrhose auftreten

[9]. Deshalb wird das Frühstadium einer Leberzirrhose oft nicht erkannt, zumal die zugrunde liegende Ursache häufig jahrelang vorliegt, bis überhaupt Beschwerden auftreten.

Das hepatozelluläre Karzinom selbst verursacht meist keine spezifischen Symptome [25]. Gelegentlich können bei randständiger Lage des Tumors frühzeitig oder als spätes Zeichen der Ausdehnung ein viszeraler Kapselschmerz oder bei Verlegung des Gallengangs ein Ikterus auftreten [25]. Außerdem ist auf allgemeine karzinomtypische Symptome wie Fieber, Nachtschweiß oder Tumorkachexie zu achten (sog. B-Symptomatik) [9]. Patienten mit Leberzirrhose als fakultativer Präkanzerose sollten regelmäßige Screeningtermine wahrnehmen, um das Entstehen von Tumoren in der Leber möglichst früh zu entdecken.

Abhängig von der Metastasierung im späteren Stadium können die unterschiedlichsten Symptome dazu kommen, je nachdem welches Organ befallen ist. So können Knochenschmerzen [9] oder Pleuraergüsse erste Zeichen einer HCC-Erkrankung sein.

Bei Patienten ohne zugrunde liegende Zirrhose ist die Symptomatik meist etwas anders. Der Tumor kann hier weitestgehend ungestört wachsen [9] und so stellen sich diese Patienten erstmals v.a. mit B-Symptomatik als Zeichen eines fortgeschrittenen Tumorstadiums vor.

Zusätzlich können die Patienten erstmalig auch durch die Symptome von krankheitsbedingten Komplikationen auffällig werden. Neben allgemeinen Komplikationen von Krebserkrankungen wie erhöhtem Thromboserisiko oder Gewichtsverlust sind auch paraneoplastische Symptome im Rahmen des HCC beschrieben worden, wie z.B. Hypoglykämie oder Hyperkalzämie [9], [25]. Außerdem wurde gezeigt, dass wässrige Durchfälle bei Leberzirrhose des Öfteren vorkommen [9], [25], bei der Kombination von Zirrhose und HCC sogar häufiger als bei Zirrhose allein [9].

Eine seltenere, aber lebensgefährlich verlaufende Komplikation ist die Tumorruptur. Hierbei kommt es zur massiven Blutung in den Bauchraum, die Patienten haben plötzlich einsetzende stärkste Bauchschmerzen sowie Abwehrspannung und Hypotension [8], [9], [50]. Die Inzidenz von spontanen HCC-Rupturen liegt etwa

zwischen 3 % und 15 % [41]. Sie sind heutzutage durch die verbesserte Früherkennung von HCC-Tumoren selten. Die Mortalitätsrate ist mit bis zu 75% sehr hoch [46].

## **1.5 Prävention**

### **1.5.1 Primärprävention**

Die Primärprävention hat die Vermeidung der Entstehung einer Erkrankung als Ziel und somit kann die Vermeidung aller auslösenden Faktoren als eine solche angesehen werden [45]. Auf das HCC bezogen wäre das z.B. der komplette Verzicht auf Alkohol oder eine nur geringe Aufnahme dessen sowie der Schutz vor einer Infektion mit Hepatitis B oder C [19]. Z.B. kann durch eine Hepatitis-B-Impfung der Infektion vorgebeugt werden, was weltweit zu einem Rückgang an HBV-Infektionen geführt hat [55], [69]. Ob der Schutz vor Aflatoxin-B1-Exposition wirklich zu einer geringeren HCC-Inzidenz führen kann, wird diskutiert [55]. Auch die Lebertransplantation als solche kann bei Nicht-HCC-Patienten als Primärprävention angesehen werden [69].

### **1.5.2 Sekundär- und Tertiärprävention**

Als Sekundärprävention werden all jene Maßnahmen bezeichnet, die das chronische Stadium der Leberentzündung bzw. -infektion therapieren und somit versuchen das Fortschreiten in die Entwicklung eines HCCs zu verhindern [45]. So kann z.B. die frühzeitige medikamentöse Therapie von Hepatitis-B- bzw. Hepatitis-C-Infektionen als Sekundärprävention für das HCC angesehen werden. Wurde früher hauptsächlich mit Interferonen therapiert [43], so erscheint v.a. bei der Virushepatitis Typ C die Therapie mit neuen, direkt antiviral agierenden Medikamenten wie z.B. Daclatasvir oder Asunaprevir vielversprechend [23], [52]. Für die Therapie mit oralen antiviralen Therapeutika bei Hepatitis B zeigen sich widersprüchliche Ergebnisse im Bezug auf die Inzidenz des HCC. So fanden *Cho et al.* einen Anstieg der HCC-Erkrankungen unter oraler antiviraler Therapie [17], wohingegen andere Autoren eine deutlich niedrigere HCC-Inzidenz bei diesen Patienten fanden [58], [77]. Laut *Colombo et al. (2014)* zeigt sich in mehreren Studien aus neuester Zeit keine Reduktion des HCC-Risikos durch die Langzeittherapie mit oralen antiviralen Medikamenten [19].

Der Tertiärprävention kommt nach erfolgreicher Ablation oder Resektion von HCC-Tumoren eine besondere Bedeutung zu, um Rezidive zu vermeiden [19], [45].

## **1.6 Diagnostik**

### **1.6.1 Index-Test und Referenzstandard**

Der Index-Test einer diagnostischen Studie ist der Test, der auf seine diagnostische Richtigkeit untersucht werden soll und im Idealfall immer durch einen Referenzstandard kontrolliert wird. In dieser Meta-Analyse galten die Ultraschall-Untersuchung sowie die CT- und MRT-Bildgebung der Leber als Index-Tests. Die Ergebnisse dieser Tests wurden dann innerhalb der Studien durch einen Referenz-Test bestätigt oder widerlegt.

Der Referenzstandard ist die Untersuchungsmethode, mit der eine bestimmte Krankheit am besten nachgewiesen werden kann. Als Goldstandard gilt der ideale Referenzstandard, der eine zweifelsfreie und fehlerlose Diagnose möglich macht [63]. Da dies im klinischen Alltag nicht immer möglich ist und viele diagnostische Tests keinen perfekten Referenzstandard darstellen [63], werden die Begriffe *Referenzstandard* und *Goldstandard* im Allgemeinen und so auch in dieser Arbeit synonym verwendet. Typische Beispiele für Referenzstandards sind die histopathologische Bestimmung einer Gewebeprobe, der operative Eingriff, das klinische Follow-up [79] oder bestimmte Laborparameter. Der Referenzstandard sollte nicht störanfällig, also nahezu unfehlbar sein [79].

Zwar sollte der Referenzstandard bei allen Patienten des Studienkollektivs durchgeführt werden [68], jedoch kann es aus Gründen der Ethik oder der Durchführung einer Studie dazu kommen, dass nicht alle Patienten den gleichen Referenzstandard-Test durchlaufen. Dies kann zur Verfälschung der Studienergebnisse führen (siehe Kapitel 3.8 Bias).

### **1.6.2 Ultraschall (US)**

Um Patienten mit HCC möglichst frühzeitig zu erkennen, werden Risikopatienten in sogenannten Screening-Programmen regelmäßig mit einer Bildgebung untersucht. Zu



diesem Zweck eignet sich vor allem der Ultraschall, da er leicht zugänglich und kostengünstig sowie ungefährlich und nicht invasiv ist [4]. Zudem liegt die Sensitivität des US als Screening-Methode zwischen 60%-80% mit einer Spezifität von über 90% [4], [14]. Generell kommt dem Screening eine große Bedeutung zu, die Prognose zu verbessern, da durch frühzeitiges Erkennen einer HCC-Entstehung eine schnelle Therapie eingeleitet werden kann. Es wurde gezeigt, dass Risikopatienten häufiger regelmäßig an US-Screening-Programmen teilnahmen, wenn sie Mitglied einer privaten Krankenversicherung waren und eine Behandlung durch einen Spezialisten auf dem Gebiet Hepatologie bekamen [1]. Potentiell negativ kann sich das Screening auf die Psyche der Patienten auswirken, etwa durch falsch positive Ergebnisse oder die regelmäßig wiederkehrende Angst vor der Untersuchung [71]. Eine zunehmende Rate an entdeckten HCC-Tumoren  $\leq 2$  cm im Durchmesser zeigt aber auch die Nützlichkeit von Screeninguntersuchungen auf [3].

Kontrovers diskutiert wird über die geeignete Taktung der Screeningintervalle. Aufgrund der Tumorverdopplungszeit wurde das Intervall auf 6 Monate festgelegt, was auch in den meisten Kliniken so durchgeführt wird. Allerdings haben Studien gezeigt, dass es keinen Unterschied im Überleben von Patienten gab, die alle 6 bzw. 12 Monate mit Ultraschall untersucht wurden [4], [14].

Als differentialdiagnostischer Test oder für das klinische Staging ist der Ultraschall weniger geeignet. Zwar können Farbdoppler und Kontrastmittel-Verstärkung mitunter einen Hinweis auf die Vaskularisierung im Tumor und somit über die Dignität liefern, jedoch werden kleine ( $\varnothing \leq 2$  cm) oder tiefliegende Tumoren vor allem in der zirrhotischen Leber im Ultraschall nur schwer erkannt.

### **1.6.3 Computertomographie (CT)**

Die Computertomographie hat eine ähnliche Sensitivität wie der Ultraschall im Bezug auf die HCC-Detektion (60%-90%). Dennoch ist laut European Association for the Study of the Liver (EASL) und der Association for the Study of Liver Diseases (AASLD) ein Tumor  $\geq 2$  cm in der zirrhotischen Leber auch dann ohne histologische Bestätigung als HCC zu diagnostizieren, wenn dieser im CT oder im MRT mit typischen Vaskularisationsmustern (siehe Abschnitt 1.6.6 Kontrastmittel) in der arteriellen sowie in der portalvenösen Phase sichtbar ist (EASL- u. AASLD-Kriterien) [14], [42]. Für

Tumoren mit einem Durchmesser von 1-2 cm gilt, dass sie nur dann ohne Biopsie wie ein HCC behandelt werden sollen, wenn sie in zwei dynamischen Bildgebungen (CT, CEUS oder KM-verstärktes MRT) das typische Vaskularisationsmuster zeigen [14]. Typischerweise stellt sich das HCC im CT sehr heterogen dar. Hierfür können Fibrosen oder Verkalkungen innerhalb des Tumors verantwortlich sein [9], mitunter lässt sich auch eine im Vergleich zum umliegenden Gewebe verstärkt oder vermindert kontrastmittelangereicherte Kapsel erkennen [4].

Verschiedene CT-Varianten kommen als Diagnostik in Frage. Heutzutage kommen sowohl die modernen Multidetektor-CTs (MDCT), die multiphasischen CTs als auch Spiral-CTs oder mitunter die Arteriographie-CT zum Einsatz. Letztere ist wegen der hohen Kosten und ihrer hohen Invasivität allerdings eher selten [9]. Die Spiral-CTs konnten durch eine schnelle Bilderabfolge nach Kontrastmittelgabe zur weiteren Verbesserung der HCC-Diagnostik beitragen [8]. Die Multidetektor-CT-Untersuchung zeichnet sich durch hohe Aufnahmegeschwindigkeiten und die Möglichkeit zur dreidimensionalen Bildgebung aus [9]. Zur Beurteilung von Lebertumoren wird oft auf die Verwendung von Kontrastmitteln zurückgegriffen.

Mit ca. 10 mSv Strahlenbelastung pro Abdominal-CT ist das Karzinomrisiko für den Patienten durch wiederholte CT-Untersuchungen erhöht [38], sodass der Einsatz dieses Diagnoseverfahrens immer bewusst abgewägt werden sollte.

#### **1.6.4 Magnetresonanztomographie (MRT)**

Die höchste Sensitivität und Spezifität aller Bildgebungen hat die Magnetresonanztomographie. Sie ist mit 90%-100% [4] vergleichbar mit der des multiphasischen CT. Auch ist bei der MRT-Untersuchung keine Strahlenbelastung gegeben. Durch diese Vorteile ist sie in vielen Kliniken weltweit als Standard für die HCC-Diagnostik angesehen [8], [9]. Allerdings ist die MRT als solche immer noch sehr teuer und in vielen, vor allem kleineren Einrichtungen nicht vorhanden. Deshalb ist die MRT als Screeningmethode ungeeignet, zur genauen Charakterisierung und für das Tumor-Staging aber unverzichtbar [4].

Die Diffusions-gewichtete MRT-Untersuchung (Diffusion-weighted imaging, DWI) macht Unterschiede in der Bewegung von Wassermolekülen im Gewebe sichtbar [4].

So wird Tumorgewebe innerhalb der Leber durch hohe Signalstärke gut sichtbar – jedoch lässt sich dadurch nicht zwischen malignem und benignem Tumor unterscheiden [10]. Laut *Bolog et al. (2011)* ist die DWI v.a. für die Entdeckung kleinerer Knoten in der Leber hilfreich, sollte aber immer in Zusammenschau mit T1- und T2- gewichteten unverstärkten und Kontrastmittel-verstärkten Bildgebungen ausgewertet werden [10].

### **1.6.5 Alpha-Fetoprotein (AFP)**

Als weltweit wichtigster Tumormarker und Verlaufsparemeter wird im Serum die Menge an  $\alpha$ -Fetoprotein gemessen. Dieses Glykoprotein wird im Fetus im Dottersack und der Leber gebildet und sinkt im Laufe des ersten Lebensjahres physiologischer Weise ab. Ein fehlendes Absinken oder ein Wiederanstieg der Werte deuten somit auf eine Schädigung der Leber hin [9]. Der Normwert eines Erwachsenen liegt bei etwa 20 ng/ml. Werte zwischen 50 und 100 ng/ml gelten als malignitätsverdächtig und sollten abgeklärt werden. Bei AFP-Werten über 100 ng/ml mit einer entsprechenden positiven Bildgebung kann die Diagnose eines HCC gestellt werden [69]. Allerdings können auch andere Erkrankungen zu einer Erhöhung des AFP-Wertes führen, wie z.B. Hepatitis oder Leberzirrhose oder tumoröse Erkrankungen des Pankreas oder der Gallenblase [9] sowie metastasierte Tumoren anderen Ursprungs, was die falsch positiven Vorhersagewert des AFPs erhöht [71]. Die Sensitivität des AFP-Wertes für die HCC-Diagnostik liegt – abhängig vom Cut-off-Wert – zwischen 40% und 60%, seine Spezifität ist hingegen besser mit etwa 75%-90% [25], [69], [M44]. Die Sensitivität nimmt mit zunehmend hohem Cut-off-Wert ab, wohingegen die Spezifität dann steigt [M44]. Ein normaler AFP-Wert schließt das Vorliegen eines Tumors nicht aus, die AFP-Messung ist somit für das Screening für HCC nicht sinnvoll [4]. Der positiv prädiktive Wert des AFPs liegt mit 10-30% [8] recht niedrig. Außerdem steht die Höhe des AFP-Wertes nicht in direkter Wechselwirkung mit Größe des Tumors, Tumorstadium oder Prognose [25] und ist somit immer nur als Teil einer Diagnostik zu sehen und niemals allein aussagekräftig.

### **1.6.6 Kontrastmittel**

Um das in der Regel sehr gut durchblutete HCC in der Bildgebung sichtbar zu machen, werden Kontrastmittel benötigt. Hierbei unterscheidet man extrazelluläre

Kontrastmittel (alle CT-Kontrastmittel und der Großteil der MRT-Kontrastmittel), die aus dem Gefäß in den extrazellulären Raum diffundieren, von hepatobiliären Kontrastmitteln (ein kleiner Teil der MRT-Kontrastmittel), welche in die Hepatozyten aufgenommen werden [64]. Die kontrastmittelunterstützte Bildgebung findet bei CT- und MRT-Untersuchungen seit langem Verwendung, ist für den Ultraschall der Leber jedoch relativ neu. Hier werden keine Flüssigkeiten im herkömmlichen Sinne appliziert, sondern ein flüssiges Gemisch, welches Mikrobläschen aus verschiedenen Gasen enthält [64]. Im Falle der frühen Kontrastmittel wie z.B. Levovist® ist dieses Gas Luft, bei neueren Mitteln wie SonoVue® oder Optison® handelt es sich um schwer wasserlösliche Gase wie Schwefelhexafluorid bzw. Perfluoropropan [25], [36], [67]. Die Mikrobläschen dieser Kontrastmittel werden bei hohen Schalldrücken in nicht-lineare Schwingung versetzt, sodass sie sich gut vom umgebenden Gewebe unterscheiden [48] und somit z.B. die Blutversorgung eines Gewebes gut dargestellt werden kann. So lassen sich in der Leber mit diesen Kontrastmitteln in Real-Time-Einstellung alle Phasen der Durchblutung darstellen [25], [36]: Die arterielle, die portalvenöse und die späte Phase. So wird eine Hypervaskularisierung in der arteriellen Phase [11], [25] und ein beschleunigtes sogenanntes „Washout“ (Ausschwemmen des Kontrastmittels) während der portalvenösen Phase [4], [25], [67] als HCC-typisch angesehen. Erstere wird bereits nach 10-20 Sekunden, letztere nach 30-45 Sekunden sichtbar [25], [67]. Die Spätphase, ca. 2 Minuten nach Gabe des Kontrastmittels, dient zur Beurteilung des Parenchyms und eventuell vorhandenen Metastasen [67]. Anders als die Kontrastmittel für CT und MRT werden diese neuartigen US-Kontrastmittel nicht über die Niere ausgeschieden. Somit sind Nierenfunktionstests vor der Untersuchung nicht nötig, außerdem sind die Ultraschall-Kontrastmittel besonders gut verträglich und es kommt nur äußerst selten, in ca. 0,0086% der Fälle, zu schwerwiegenden Nebenwirkungen [48]. Die neuen Ultraschalluntersuchungen mit Mikrobläschen-Kontrastmittel zeigen vielversprechende Ergebnisse, z.T. sogar ähnlich der diagnostischen Richtigkeit von MRT-Untersuchungen [24].

Auch in der CT-Bildgebung hat die Verwendung von Kontrastmitteln eine große Bedeutung. Schnell nacheinander gewonnene Bilder zu unterschiedlichen Zeitpunkten nach Kontrastmittelgabe - das multiphasische CT - verschafft einen Überblick über die Durchblutung des Organs und des Tumors [9]. Im sogenannten triphasischen CT kann

die Durchblutung des Organs vor Kontrastmittelgabe, in der arteriellen Phase und in der portalvenösen Phase genau beobachtet werden [8], [9].

Für die MRT-Untersuchung haben sich in besonderem Maße Gadobenat-Dimeglumin, Gadoxetsäure, Ferucarbotran und superparamagnetische Eisenoxid-Partikel als geeignete Kontrastmittel gezeigt. Letztere haben als Zielzellen die Kupfferzellen der Leber [4]. Die Kombination aus superparamagnetischen Eisenoxid-Partikeln und Gadolinium-haltigen Kontrastmitteln ist in Bezug auf die diagnostische Richtigkeit der MRT-Untersuchung mit der der CT-Arteriographie vergleichbar [9].

#### **1.6.7 Feinnadelbiopsie**

Der Goldstandard in der Diagnostik des HCC ist nach wie vor die Biopsie. Sie sollte allerdings nur dann erfolgen, wenn sich hieraus für den Patienten eine therapeutische Konsequenz ergibt [35]. Ein Vorteil der Biopsie ist, dass das Gewebe auf sein biologisches Verhalten getestet werden kann. Somit können molekulare Expressionsmuster bestimmter Rezeptoren ermittelt und eventuell eine gezielte Krebstherapie mit neuen Medikamenten durchgeführt werden [42]. Bei eindeutig malignitätstypischem Kontrastmittelverhalten in der Bildgebung bei Risikopatienten kann auf eine Feinnadelbiopsie verzichtet werden. Ebenso sollte bei Rundherden, die kein charakteristisches Kontrastmittelverhalten in der ersten Bildgebung zeigen und die einen mindestens 2 cm großen Durchmesser haben, eine zweite kontrastmittelunterstützte Bildgebungsmethode gewählt werden. Eine auffällige zweite Bildgebung würde eine Gewebeentnahme unnötig machen. Bei Rundherden, die kleiner als 1 cm im Durchmesser sind, sollte eine eventuelle Größenzunahme nach 3 Monaten z.B. durch Ultraschall kontrolliert werden und bei Größenprogredienz erneut durch eine kontrastmittelgesteuerte Bildgebung auf auffällige Darstellung untersucht werden. Befunde zwischen 1 und 2 cm Durchmesser sollten dann biopsiert werden, wenn ein guter Zugangsweg besteht und ansonsten mittels einer zweiten kontrastmittelunterstützten Bildgebung auf Malignitätszeichen untersucht werden. Fehlen diese, erfolgt wie bei Befunden < 1 cm im Durchmesser, die regelmäßige Kontrollbildgebung alle 3 Monate [35].

## **1.7 Staging**

Wie weit ein Tumor fortgeschritten ist beeinflusst allgemein die Prognose und die Therapieoptionen für den Patienten. Bei Lebertumoren spielt außerdem eine Rolle, dass mit Fortschreiten des Tumors auch die Funktion dieses lebenswichtigen Organs beeinträchtigt werden kann und somit die Prognose weiter verschlechtert wird. Insofern ist das sogenannte Staging ein wichtiger Bestandteil der Behandlung [8], [14]. Über die Jahre haben sich verschiedene Staging-Systeme entwickelt, die alle mehr oder weniger kontrovers diskutiert werden. Noch gibt es keinen Goldstandard und theoretisch bleibt es jeder Klinik selbst überlassen, nach welchem System sie vorgeht. Allerdings gibt es in den verschiedenen Ländern eindeutige Trends zu einem bestimmten Staging-Verfahren. Wichtig ist, dass die unterschiedlichen Patientenuntergruppen beachtet werden, da sich je nach Größe und Ausbreitung des Tumors verschiedene Therapieoptionen bieten [8], [14]. Die Patientenuntergruppen sollten dabei nach ihrem voraussichtlichen Überleben eingeteilt werden: innerhalb einer Gruppe sollte das Überleben einheitlich sein und sich gegenüber einer anderen Patientenuntergruppe deutlich unterscheiden [8]. Im Folgenden werden die verschiedenen Staging-Systeme kurz vorgestellt. Bis auf das TNM-Staging-System beziehen sie sich alle auf HCC bei Leberzirrhose.

### **1.7.1 TNM-Staging-System**

Das klassische TNM-Staging-System gibt es nicht nur für das hepatozelluläre Karzinom, sondern für viele Tumoren unterschiedlichster Art. Hierbei gehen die Größe des Primärtumors und eventuelle Invasion der Nachbargewebe (T = Tumor), die Anzahl der Lymphknotenmetastasen (N = Nodules) und die Anzahl der Fernmetastasen (M= Metastases) in die Berechnung des Staging mit ein. Es ist vor allem für chirurgische Eingriffe von Bedeutung und somit weit verbreitet [6]. Allerdings berücksichtigt diese Einteilung nicht die Leberfunktion, das Tumorstadium oder den körperlichen Zustand des Patienten, sodass diese Parameter zusätzlich erhoben werden müssen, um die individuelle Prognose von HCC-Patienten einschätzen zu können [35].

### **1.7.2 Okuda-Staging-System**

Dieses Staging-System, welches außer der Tumorausbreitung auch die Leberfunktion in Betracht zieht [14], war das erste für das HCC entwickelte Staging-System [6]. Es fand daher lange Zeit intensive Verwendung weltweit [6], [8]. In die Einteilung der Stadien gehen Parameter wie Aszites, Bilirubin und Albuminwerte ebenso ein wie die Tumorgöße, sodass insgesamt zwischen 0 Punkte bei niedrigem Risikoprofil (Okuda Stadium I) und 4 Punkte bei hohem Risiko (Okuda Stadium III) erreicht werden können. Allerdings ist dieses Staging-System nur hilfreich, um eine Untergruppe mit schlechter Prognose, also mit weit fortgeschrittenem Tumorleiden zu identifizieren [6], [8], [14]. Diese Patienten des Okuda-Stadiums-III sollten nur noch palliativ behandelt werden [8]. Dank der heutigen Screening-Programme werden die Tumoren immer früher entdeckt und die Patienten befinden sich daher in einem frühen Stadium. Um die prognostischen Unterschiede und bestmöglichen Therapien aufzudecken, eignet sich das Okuda-Staging-System nicht. Es findet daher heute nicht mehr die breite Anwendung, die es noch vor 10 Jahren hatte.

### **1.7.3 Cancer of the Liver Italian Program (CLIP) -Staging- System**

Dem Okuda-System überlegen zeigte sich das CLIP-Staging-System [8]. Dieses System berücksichtigt sowohl die Tumormorphologie als auch die Tumorgöße, sowie AFP-Werte im Serum und Beteiligung der Portalvene ebenso wie die Child-Pugh-Klasse (siehe Abschnitt 1.8.2 Lebertransplantation). Dadurch lassen sich gerade in den Okuda Stadien I (Patient ohne Aszites, Serumalbumin  $>3\text{mg/dl}$ , Serumbilirubin  $<3\text{mg/dl}$  und geringe Tumorgöße) und Okuda Stadium II (Patient mit mindestens einem Risikofaktor wie z.B. Serumbilirubin  $>3\text{mg/dl}$ ) Unterschiede finden, die für die Prognose und Therapie des Patienten von Bedeutung sind.

Das CLIP-Staging-System wird kritisiert für seine unzureichend exakte Therapieempfehlung in den unterschiedlichen Stadien sowie für eine zu ungenaue Überlebensvorhersage in unterschiedlichen Patientengruppen [6]. Auch scheint das CLIP-System eine Schwäche in Bezug auf Patientengruppen zu haben, die sich radikalen Therapiemaßnahmen wie Transplantation oder Resektion bereits unterzogen haben [8].

Trotzdem wurde in einer Studie aus dem Jahr 2010 im Vergleich zwischen TNM, Okuda, CLIP, BCLC und JIS festgestellt, dass das CLIP-System die am besten geeignete Modell zur Langzeit-Prognose bei Patienten mit HCC darstellt [39].

#### **1.7.4 Japanese Integrated Staging Score (JIS)**

Dieses Staging-System berücksichtigt sowohl Child-Pugh-Klasse als auch die TNM-Klassifikation. Im japanischen Klinikalltag wird fast ausschließlich dieser Staging-Score verwendet. Außerhalb Japans findet er allerdings kaum Beachtung, da unklar ist, inwiefern sich das System, das an großen japanischen Patientengruppen getestet wurde, auch auf andere Bevölkerungsgruppen anwenden lässt [55].

#### **1.7.5 Barcelona Clinic for Liver Cancer-Staging-System (BCLC)**

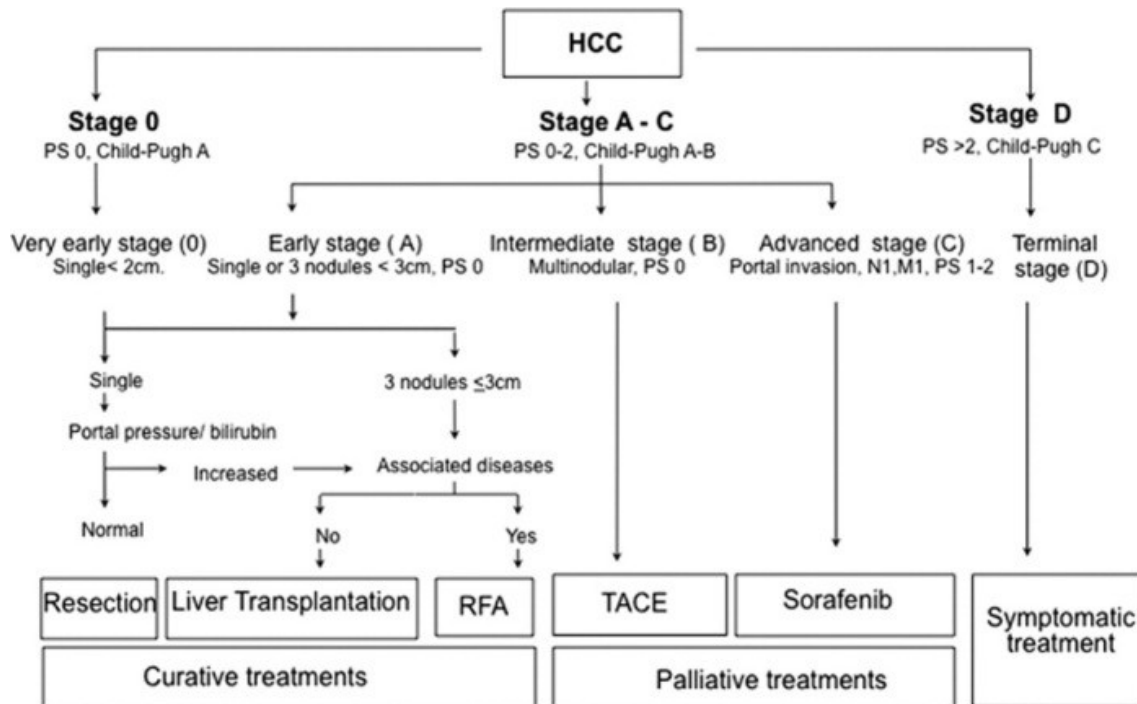
Das heutzutage meist verwandte Staging-System ist das BCLC [30]. Neben den Parametern Tumorgröße und Leberfunktion spielt hier auch das klinische Bild des Patienten eine Rolle. Nach WHO-Kriterien wird das Aktivitätsniveau des Patienten in die Stadien 0 (Volle Aktivität, keine Symptome) bis 4 (Komplett inaktiv, an Bett oder Stuhl gebunden) eingeteilt. Ein Patient im frühen Krankheitsstadium mit kleinem singulärem Leberknoten und hoher körperlicher Leistungsfähigkeit erhält das Stadium 0, wohingegen Patienten mit größeren oder multiplen Leberknoten bzw. bei Portalinvasion oder niedrigerer körperlicher Leistungsfähigkeit die Stadien A-C beschreiben. Patienten im Terminalstadium der Erkrankung mit stark eingeschränktem körperlichem Aktivitätsniveau bilden schließlich das Stadium D (siehe Abb. 1).

Besonders vorteilhaft am BCLC-Staging-System ist, dass es das Tumorstadium mit der am besten passenden Therapie und einer möglichst genauen Lebenserwartung verknüpft, sodass sehr individuell geplant und behandelt werden kann [6], [8]. Es ist vor allem hilfreich für das Staging im Frühstadium der Erkrankung und kann hier helfen, zwischen Patienten zu unterscheiden, die von aggressiver Therapie wie z.B. einer Resektion profitieren würden und solchen, die besser mit lokal ablativen Methoden behandelt werden sollten [8].

Anfangs war die fehlende externe Bestätigung dieses Staging-Verfahrens ein Manko des BCLC-Systems [8]. Trotzdem wurde es immer weiter aktualisiert und wurde sowohl von der European Association for the Study of the Liver (EASL) als auch von der



American Association for the Study of Liver Diseases (AASLD) akzeptiert [6]. Mittlerweile wurde das BCLC-Staging-System sowohl in Europa als auch in Japan validiert und kann somit heutzutage als das am besten geeignete Staging-System für HCC-Patienten angesehen werden [31].



**Abb. 1:** Das BCLC-Staging-System für das HCC mit daraus resultierender Therapieempfehlung [15].

M=Metastasis Classification; N=Node Classification; PS=Performance Status;

RFA=Radiofrequenzablation; TACE=Transarterielle Chemoembolisation

## 1.8 Therapie

Die verschiedenen Therapieoptionen für das hepatozelluläre Karzinom unterscheiden sich deutlich in ihrer Effektivität und Zielsetzung. Da die Erkrankung in mehreren Stadien verläuft, hängt es unter anderem stark vom Krankheitsstadium ab, welche Therapie gewählt wird. Insbesondere deshalb ist eine möglichst frühzeitige und exakte Diagnostik von enormer Bedeutung.

In frühen Stadien ist die kurative Therapie meist das Ziel der Behandlung. In späten Stadien versucht man, möglichen Rezidiven vorzubeugen oder das weitere Fortschreiten der Erkrankung (Tumorausbreitung, Metastasierung) zu verhindern. Wenn das Malignom bereits so weit fortgeschritten ist, dass die Leberfunktion stark eingeschränkt ist und der Patient keine Aussicht auf eine erfolgreiche Transplantation

des Organs hat, bleibt oft nur die palliative Therapie mit den Zielen der Schmerzlinderung und des Erhalts der Lebensqualität.

Um die am besten geeignete Therapie auswählen zu können, werden die Tumorgröße, die Leberfunktion, der Allgemeinzustand des Patienten mit eventuellen Komorbiditäten und die geschätzte Überlebenswahrscheinlichkeit bewertet.

### **1.8.1 Tumorresektion**

Der Goldstandard der kurativen Therapie ist immer noch die chirurgische Intervention in Form einer Leberresektion, insbesondere für Tumoren kleiner als 2 cm im Durchmesser [30]. Durch die Tumorresektion besteht theoretisch eine sehr gute Heilungschance – vorausgesetzt, der Tumor befindet sich noch im Frühstadium, ist somit nicht metastasiert und unifokal. Denn dann ist die Wahrscheinlichkeit am größten, dass der Tumor im Ganzen entfernt werden kann.

Das Hauptproblem der Resektion liegt in der Rezidivrate. Es wird hier unterschieden zwischen wirklichen Rezidiven, die kleinsten Metastasen in der Leber an derselben Stelle des Primärtumors entsprechen, und neu entstehenden Tumoren [31]. Mit 60-70% am häufigsten sind die wirklichen Rezidive [14], [31]. Eine erneute Resektion ist nur in wenigen Fällen möglich, da durch die erste Operation eine Streuung des Tumors stattfindet und die Rezidive somit meist multifokal auftreten [14].

### **1.8.2 Lebertransplantation**

Der Ersatz des vom Tumor befallenen Organs bietet einerseits den Vorteil, dass intrahepatische Metastasen oder Gefäßinvasionen gleich mit entfernt werden, andererseits wird auch eine zugrunde liegende Zirrhose geheilt [30]. Wie bei allen Transplantationen ist eins der größten Probleme jedoch die begrenzte Anzahl an Organen. Dies macht die genaue Auswahl der Patienten so entscheidend, die am besten von einer Transplantation profitieren würden. Daher eignet sich die Lebertransplantation vor allem für Patienten im frühen Stadium, deren solitäre Tumoren kleiner als 5 cm im Durchmesser sind oder für Patienten, die maximal 3 Tumoren mit je höchstens 3 cm im Durchmesser haben [35]. Diese sogenannten Mailand-Kriterien haben sich als prognostische Faktoren für ein sehr gutes postoperatives Outcome bewährt [14]. Die 5-Jahres-Überlebensrate dieser Patienten

beträgt über 70 % mit einer geringen Rezidivrate [14], [31], [53]. Der potentielle Gewinn für den Patienten und das Risiko des Eingriffes sollten bei der für die Transplantation infrage kommenden Patienten sorgfältig gegeneinander abgewogen werden [65].

Der für die Einteilung und Vorhersage des klinischen Outcomes nach Lebertransplantation am häufigsten benutzte Score ist der Child-Pugh Score (CP) (siehe Abb. 2) [7]. Seine einfache Anwendung machte ihn im klinischen Alltag beliebt und so wird er auch heute noch bei Patienten mit Zirrhose regelhaft bestimmt, um eine Aussage über das Stadium der Erkrankung machen zu können. Zur Bestimmung dieses Scores gehen folgende Parameter ein: Serumbilirubin und Serumalbumin, der Grad an Aszites, eine eventuell vorliegende Enzephalopathie und die Thromboplastinzeit als International Normalized Ratio (INR), d.h. als standardisierter Quick-Wert [30], [55].

<b>Table 1: Scoring* Child–Pugh–Turcotte criteria for end-stage chronic liver disease</b>			
Parameter	1 point	2 points	3 points
Encephalopathy grade	None	1 or 2	3 or 4
Ascites	Absent	Slight	Moderate
INR	<1.7	1.7–2.3	>2.3
Total bilirubin, µmol/L	<17	17–51	>51
Albumin, g/L	>35	28–35	<28
* Classification is derived by summing points for each of the 5 parameters: Class A, 5–6 points; Class B, 7–9 points; and Class C, 10–15 points. INR = international normalized ratio.			

**Abb. 2:** Der Child-Pugh-Score [7].

Ein Patient mit Child A gilt als gut kompensiert und weist in den 5 Parametern normale oder fast normale Werte auf, wohingegen ein Child C-Patient eine schwere Beeinträchtigung aufweist [55].

Eine Möglichkeit zur besseren Ausschöpfung der Organknappheit ist die Splitliver-Transplantation. Diese v.a. in der Pädiatrie häufig angewandte Therapieoption findet zunehmend auch bei Erwachsenen Verwendung. Hierbei wird ein Spenderorgan in zwei Teile aufgeteilt und zwei Organempfängern implantiert. Dadurch kann eine

Leberspende für mehrere Patienten und das Angebot an Leichentransplantaten effektiver genutzt werden [18].

Zusätzlich gibt es die Möglichkeit, Patienten auf der Warteliste bereits chirurgisch oder medikamentös zu behandeln, um zu verhindern, dass der Tumor weiter wächst und die Mailand-Kriterien somit nicht mehr erfüllt werden (sog. „Bridging“) . Hierfür eignen sich zum Beispiel die Resektion, aber auch lokal ablativ Verfahren wie Radiofrequenzablation oder die Chemoembolisation [14], [30], [31], [35].

### **1.8.3 Radiofrequenzablation (RFA)**

Ist ein Patient z.B. aufgrund seiner Komorbiditäten kein Kandidat für einen chirurgischen Eingriff, so gilt als vielversprechende Therapieoption die Radiofrequenzablation. Auch sie hat einen kurativen Ansatz und wird somit v.a. im sehr frühen und frühen Erkrankungsstadium (BCLC Stage 0-A) angewandt [15]. In der aktuellen S3-Leitlinie *Diagnostik und Therapie des hepatozellulären Karzinoms Version 1.0* [35] vom Mai 2013 wird die RFA als Goldstandard der perkutanen Lokalthherapie vorgestellt. Bei der RFA wird in den Tumor eine Nadel eingeführt, welche einen elektrischen Wechselstrom mit einer Frequenz zwischen ca. 400 und 500 kHz abgibt [20], [31], wodurch das Tumorgewebe erhitzt und somit zerstört wird. Bei einer gleichzeitig unveränderten Anzahl an operativen Therapien wurde in den USA zwischen den Jahren 2000 und 2010 eine Zunahme der Ablationstherapien v.a. bei kleinen Tumoren beobachtet [75]. Obwohl die RFA in Bezug auf Rezidivrate ähnliche oder gleiche Ergebnisse liefert wie die Resektion [20], hat sie zum Nachteil, dass die technische Ausrüstung für die RFA sehr teuer ist und es kommt relativ häufig zu unerwünschten Ereignissen wie Blutungen oder Pleuraergüssen [14], [31]. Die Komplikationsrate liegt hier bei ca. 8,9%, die Mortalität bei 0,5% [31].

### **1.8.4 Transarterielle Chemoembolisation (TACE)**

Als Therapieoption für das weiter fortgeschrittene Stadium (multinodales HCC, BCLC Stage B) steht die transarterielle Chemoembolisation zur Verfügung. Diese wurde aufgrund der Beobachtung entwickelt, dass das HCC seine Blutversorgung fast ausschließlich aus Ästen der Arteria hepatica gewinnt, sodass eine Unterbrechung dieser Blutzufuhr zum Untergang des kanzerösen Gewebes führen muss [31]. Nach

Gabe eines Chemotherapeutikums, welche möglichst sehr gezielt eingesetzt werden sollte, um das umgebende gesunde Gewebe zu schonen, wird die arterielle Blutzufuhr der Leber unterbunden. Hierfür werden Gelfoam®-Würfel [14], [31] oder Medikamenten-freisetzende Mikrosphären benutzt [16], [56]. Außerdem wird nach antiangiogenetischen Mitteln gesucht, die die Effektivität von TACE verbessern könnten [44].

Da dieses Verfahren zwar das weitere Wachstum des Tumors verhindern oder verzögern kann, jedoch keine Heilung bringt, findet diese Therapie vor allem in der Palliativversorgung von HCC-Patienten Verwendung.

### **1.8.5 Systemische Therapie**

Eine weitere Behandlungsoption für Palliativpatienten ist die Chemotherapie. Da das HCC relativ unsensibel für Chemotherapie ist und die Patienten durch die eingeschränkte Leberfunktion entsprechend schlecht auf diese Mittel ansprechen und nur eine angepasste Dosierung vertragen, ist die Indikationsstellung vorher genau zu prüfen [8], [31]. Für Patienten mit weit fortgeschrittenem Tumorstadium zeigte sich v.a. die Behandlung mit Sorafenib als effektiv [51]. Dieser Multikinase-Inhibitor wirkt nachweislich lebensverlängernd [51] und ist aufgrund der im Vergleich weniger schwerwiegenden Nebenwirkungen wie Durchfall und Hand-Fuß-Ausschlag zur Therapie der ersten Wahl für HCC-Patienten geworden, die nicht für eine effektivere Behandlung in Frage kommen [15]. Eine vorangehende Therapie mit TACE vor der Sorafenibbehandlung zeigt positive Ergebnisse in Bezug auf das Gesamtüberleben [56], [70].

### **1.8.6 Symptomatische Therapie**

Im Endstadium der Erkrankung (Child-Pugh C, Aktivitätsniveau des Patienten >2, BCLC Stage D) kann dem Patienten noch durch eine symptomatische Therapie geholfen werden. Das Fortschreiten des Tumors ist nicht aufzuhalten, im Vordergrund stehen das Befinden des Patienten und die Linderung von Schmerzen sowie der Erhalt bzw. die Verbesserung von Lebensqualität. Im Rahmen dieser Ziele kommen Schmerztherapien unterschiedlichster Art und andere Medikamente, die dem

Patienten Linderung verschaffen (Laxantien, Sauerstoff, Beruhigungsmittel), zum Einsatz. Weniger als 10 % dieser Patienten überleben länger als ein Jahr [31].

## **2. FRAGESTELLUNG UND ZIELSETZUNG DIESER ARBEIT**

Das hepatozelluläre Karzinom ist eine der häufigsten bösartigen Krankheiten weltweit. Da seine insgesamt schlechte Prognose oft nur durch eine frühzeitige Erkennung der Erkrankung gebessert werden kann, spielt die Diagnosestellung eine wichtige Rolle. Diese ist aber weiterhin dadurch erschwert, dass geeignete bildgebende Verfahren nicht überall verfügbar sind. Die weiterführende Diagnostik soll möglichst wenig invasiv sein und eine genaue Aussage liefern. Welches Diagnoseinstrument sich dafür am besten eignet, soll in dieser Arbeit herausgefunden werden.

Ziel dieser Dissertation ist es, die aktuelle Literatur zu dieser Fragestellung systematisch nach geeigneten Studien zu durchsuchen, die ausgewählten Artikel hinsichtlich ihrer Qualität und Durchführung zu bewerten und quantitativ in Form einer Metaanalyse zusammenzufassen.

Bei der Bewertung wird besonderes Augenmerk auf einen geeigneten Referenzstandard (siehe Kapitel 1.6.1, S.9) und seine Unabhängigkeit vom Index-Test, auf die Ein- und Ausschlusskriterien der Studienpopulation und auf ein ausreichendes Follow-up gelegt. Demnach werden die einzelnen Studien mit einem Evidenzgrad versehen und unter dessen Berücksichtigung in der Metaanalyse zusammengefasst.

Außerdem werden die Impaktfaktoren der Fachzeitschriften, in welcher die Artikel veröffentlicht wurden, in die Auswertung mit einbezogen. So soll untersucht werden, ob es einen Zusammenhang gibt zwischen der Qualität der Fachzeitschrift und dem Niveau des Originalartikels in Bezug auf dessen methodisches Vorgehen.

Als Ziel dieser Arbeit wird die diagnostische Genauigkeit von Ultraschall, CT und MRT in der Erkennung des HCC berechnet und bewertet und es wird ein Überblick über die derzeit vorhandene Literatur zum Thema gegeben.

## 3. MATERIAL UND METHODEN

### 3.1 Basisliteratursuche

Im November 2011 wurde die Arbeit an dieser Meta-Analyse begonnen. Bereits im Frühjahr desselben Jahres wurde eine systematische, umfassende Suche zum Thema Diagnostik des hepatozellulären Karzinoms im Rahmen der Recherche zum „Leitlinienreport der S3-Leitlinie Diagnostik und Therapie des hepatozellulären Karzinoms“ durchgeführt. Die hierfür als relevant empfundenen Studien wurden als erste darauf untersucht, ob sie für diese Meta-Analyse geeignet sind. Der Suchvorgang für die in der Leitlinie verwendeten Studien sei hier noch einmal kurz aufgeführt: In der Datenbank PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) wurde unter Beachtung der **Limits: Humans, Clinical Trial, Meta-Analysis, Practice Guideline, Randomized Controlled Trial, Review, English, German Publication date from 2000/01/01 to 2011/03/31** eine Suchanfrage mit den MeSH-Begriffen **“Carcinoma, Hepatocellular”[MeSH] AND diagnosis[MeSH] AND (“ultrasonography”[Subheading] OR “ultrasonography”[All Fields] OR “ultrasound”[All Fields] OR “ultrasonography”[MeSH Terms] OR “ultrasound”[All Fields] OR “ultrasonics”[MeSH Terms] OR “ultrasonics”[All Fields]) OR (“computers”[MeSH Terms] OR “computers”[All Fields] OR “computer”[All Fields]) AND (“tomography, x-ray computed”[MeSH Terms] OR (“tomography”[All Fields] AND “x-ray”[All Fields] AND “computed”[All Fields]) OR “x-ray computed tomography”[All Fields] OR “tomography”[All Fields] OR “tomography”[MeSH Terms])) OR MRT[All Fields]** durchgeführt.

Als Erstausswahlkriterium dienten die Abstracts der Studien. Einschlusskriterien waren: Deutsche oder englische Publikationssprache, Diagnostik des hepatozellulären Karzinoms durch Ultraschall, CT- oder MRT-Bildgebung mit oder ohne Kontrastmittel, randomisiert kontrollierte oder klinische Studien, retrospektive oder prospektive Studien, Histopathologie oder Resektat des Tumors als Referenzstandard oder Follow-up Bildgebung mindestens 6 Monate danach (gemäß EASL Konsensus bzw. AASLD Guideline), Angabe von Sensitivität und Spezifität oder Daten, aus denen sich diese



berechnen lassen, und Datum der Veröffentlichung zwischen dem 01.01.2000 und dem 31.03.2011.

Als Ausschlusskriterien galten: Review (Übersichtsartikel), Narrative Overview, Case Report (Fallbericht), andere Sprachen als Englisch oder Deutsch, keine HCC-Patientenpopulation und sequentielle Tests, zum Ausschluss eines Selectionbias. Als sequentielle Tests galten alle diagnostischen Untersuchungen, die ein HCC entweder direkt (CT, MRT) oder indirekt (Laborparameter) aufzeigen können, bis auf die nicht Kontrastmittel-verstärkte Ultraschalluntersuchung. Als sequentieller Test gilt eine Untersuchung, die nur aufgrund des Ergebnisses einer vorangegangenen Untersuchung durchgeführt wurde. Da die Ultraschalluntersuchung Bestandteil des Routinescreenings für Risikopopulationen ist, wurde ein im Ultraschall entdeckter Knoten mit darauf folgender weiterer Diagnostik nicht als sequentieller Test gewertet. Aufgrund eines erhöhten Laborwertes oder durch CT-Bilder vermutete Leberläsionen wurden dagegen als sequentielle Tests gewertet.

Zusätzlich wurde eineinhalb Jahre später erneut eine Suche auf PubMed (<http://www.ncbi.nlm.nih.gov/pubmed> Stand: 01.05.2013) durchgeführt, um mögliche weitere Studien zum Thema zu finden. Es wurden deutsche und englische Studien gesucht, dabei wurden keine Reviews mit einbezogen. Die Suche wurde im Mai 2013 durchgeführt und es wurde ein Zeitraum von 13 Jahren eingeschlossen (Januar 2000 bis Mai 2013). Außerdem war das Vorhandensein eines Abstracts ein Einschlusskriterium, da nach diesem Abstract die Vorauswahl der Studien getroffen wurde.

Für die systematische Suche auf PubMed wurden immer die gleichen Limits verwendet. Diese waren: Randomized Controlled Trial, Clinical Trial, Meta-Analysis, Practice Guideline, Abstract available, Publication date from 2000/01/01 to 2013/05/01, Humans, English, German. Überlappungen zum ersten Suchvorgang waren möglich; mehrfach gefundene Artikel wurden zur Basisliteratursuche gezählt.

### **3.2 Diagnostische Studien und Ultraschall**

Es wurden die MeSH-Begriffe „hepatocellular carcinoma“ und „ultrasound“ in PubMed mit oben genannten Limits eingegeben (**hepatocellular carcinoma[MeSH Terms] AND**

**ultrasound[MeSH Terms]**). Durch Lesen der Abstracts wurde bestimmt, ob eine Studie ein- oder ausgeschlossen wird. Von eingeschlossenen Studien wurde der Volltext besorgt.

### **3.3 Diagnostische Studien und CT**

Für die Suche von Studien mit Computertomographie wurden die MeSH-Begriffe „hepatocellular carcinoma“ und „computed tomography“ in PubMed gesucht (**hepatocellular carcinoma[MeSH Terms] AND computed tomography[MeSH Terms]**). Die Abstracts wurden gelesen und von eingeschlossenen Studien wurde der Originaltext besorgt.

### **3.4 Diagnostische Studien und MRT**

Die MeSH-Begriffe „hepatocellular carcinoma“ und „magnetic resonance imaging“ wurden auf PubMed mit den oben genannten Limits gesucht (**hepatocellular carcinoma[MeSH Terms] AND magnetic resonance imaging[MeSH Terms]**). Das Lesen der Abstracts führte zum Ein- oder Ausschluss der Studien, eingeschlossene Studien wurden im Volltext genauer untersucht.

### **3.5 Kombinierte Suche diagnostischer Studien**

Als letztes wurde eine Suche gestartet, die alle drei diagnostischen Tests (US, CT und MRT) auf einmal finden sollte. Hierfür wurden diesmal nicht die MeSH-Begriffe gewählt, sondern die Rubrik „All Fields“ ausgewählt (**hepatocellular carcinoma AND diagnosis AND (ultrasound OR computed tomography OR MRI)**). Die Abstracts der Studien wurden gelesen und von eingeschlossenen Studien wurde der Volltext besorgt.

### **3.6 Diagnostische Studien und AFP**

Da die Messung von AFP im Serum für sich allein genommen eine relativ geringe Sensitivität für das HCC hat und immer als Teil der Gesamtdiagnostik gesehen werden muss, wurde keine eigene Suche für die Diagnostik mit AFP durchgeführt.

### **3.7 Zuteilung des Evidenzgrades**

Um die Studien in ihrer Qualität bewerten zu können, wurde jede einzelne Studie auf bestimmte Fragestellungen hin untersucht und das Ergebnis für jede Frage notiert. Am Ende ergab sich aus der Anzahl der Übereinstimmung bzw. Nichtübereinstimmung mit dem Fragebogen ein bestimmter Evidenzgrad. Der für diese Dissertation verwendete Fragebogen ist im Anhang 8.2 Anlage 2 zu finden. Er basiert auf dem vom *Oxford Centre for Evidence-based Medicine* erstellten Fragebogen und ist genau auf die wichtigen Fragestellungen zugeschnitten, die für diagnostische Studien ausschlaggebend im Hinblick auf ihre Qualität sind. Zur Auswertung der Fragebögen wurden die *Levels of Evidence* aus dem Jahr 2001 vom gleichen Institut angewandt (siehe Tabelle 1). So konnte jede Studie anhand der Kriterien, die sie erfüllt oder nicht, in einen Evidenzlevel eingeteilt werden. Insgesamt gibt es fünf solcher Levels, die eine absteigende Qualität der Studien anzeigen. Level 1a bezeichnet somit die am besten durchgeführten Studien, deren Ergebnisse valide sind und zur Meinungsbildung empfohlen werden können. Mit Level 5 werden Artikel bewertet, bei denen es sich z.B. um eine Expertenmeinung ohne wissenschaftliche Studie handelt oder um Zusammenfassungen zu einem Thema.

<b>Evidenzgrad</b>	<b>Definition</b>
<b>1A</b>	Systematisches Review von diagnostischen Studien mit Evidenzgrad 1
<b>1B</b>	Validierende Kohortenstudie mit gutem Referenzstandard
<b>2A</b>	Systematisches Review von diagnostischen Studien mit Evidenzgrad 2 und besser
<b>2B</b>	Explorative Kohortenstudie mit gutem Referenzstandard
<b>3A</b>	Systematisches Review von Studien mit Evidenzgrad 3b und besser
<b>3B</b>	Nicht-konsequente Studie; oder ohne durchgehend angewandten Referenzstandard
<b>4</b>	Fall-Kontroll Studie, schlechter oder nicht unabhängiger Referenzstandard
<b>5</b>	Expertenmeinung ohne kritische Auswertung, oder basierend auf Forschungsergebnissen im Labor, Physiologie oder „Grundprinzipien“

**Tabelle 1:** Einteilung der Evidenzgrade diagnostischer Studien [basierend auf Oxford Centre for Evidence-based Medicine [60], siehe Tabelle 18, im Anhang 8.4 Anlage 4]

Die Studien wurden von zwei Untersuchern unabhängig voneinander gelesen und mit einem Evidenzgrad versehen. Anschließend wurde das Ergebnis untereinander verglichen und bei ungleicher Bewertung wurde durch Diskussion ein Konsens gefunden.

### **3.8 Bias**

Als Bias bezeichnet man den systematischen Unterschied eines gemessenen Wertes vom wahren Wert [68]. Das bedeutet, dass der gemessene Wert sich systematisch vom wahren Wert unterscheidet und es somit bei erneuter Anwendung ständig zu einer Unter- oder Überschätzung des Wertes kommt. Der Bias ist systematisch, d.h. er kann durch Wiederholung nicht ausgeglichen werden [68]. Um ein Testergebnis verallgemeinern zu können, sollte ein geeignetes Patientenkollektiv ausgewählt werden. Als geeignet kann hier eine Patientengruppe angesehen werden, die die gesamte Bandbreite der Erkrankung widerspiegelt und somit auch das Patientenkollektiv darstellt, an dem der diagnostische Test gegenwärtig und in Zukunft angewandt wird [49]. Dies sind einerseits die tatsächlich Erkrankten und andererseits auch die Nicht-Erkrankten [79]. Zur Definition von Index-Test, Referenzstandard und Goldstandard, siehe Kapitel 1.6.1 Index-Test und Referenzstandard, S. 9.

#### **3.8.1 Spectrum Bias**

Die Unterschiede in dem Patientenkollektiv können eine Auswirkung auf die Ergebnisse des diagnostischen Tests haben. Beispielsweise ist es wahrscheinlicher ein positives Testergebnis zu bekommen bei einem Patienten, der stark ausgeprägte Symptome und eine weit fortgeschrittene Erkrankung hat, als bei einem asymptomatischen Patienten im Frühstadium der Erkrankung [68]. Ebenso können Komorbiditäten oder die Einnahme von Medikamenten dazu führen, dass ein Testergebnis eher negativ oder eher positiv ausfällt [81].

Der sogenannte *Spectrum Bias* sagt aus, dass sich das Testergebnis ändert, wenn der Test auf unterschiedliche Subgruppen von Patienten angewendet wird [81]. Die beiden oben genannten Punkte können Ursachen hierfür sein.

### **3.8.2 Selection Bias**

Ebenso kann eine diagnostische Studie durch den *Selection Bias* verfälscht werden. Wenn in einer ursprünglichen Population nicht alle Patienten mit der interessierenden Krankheit konsekutiv in die Studie mit aufgenommen werden und dies nicht zufällig passiert [49], so ist das Patientenkollektiv nicht repräsentativ. Für die interne Validität einer Studie ist die Kenntnis dieses Bias wichtig [68]. Die interne Validität gibt Aufschluss darüber, ob ein Test auch wirklich das misst, was er messen soll.

### **3.8.3 Verification Bias**

Der *Verification Bias* tritt auf, wenn in einer Studie der Test, der als Referenzstandard dient, nicht bei allen Patienten durchgeführt wird, sondern nur bei einer bestimmten Subgruppe. Da in diagnostischen Studien oft invasive Methoden als Goldstandard/Referenzstandard verwendet werden, werden teilweise nur die Patienten mit diesem Test nachuntersucht, die vorher im zu untersuchenden diagnostischen Test schon ein positives Ergebnis hatten [49], [68]. Ist dies der Fall so spricht man vom Partial Verification Bias. Wird bei den übrigen Patienten dann ein anderer Test durchgeführt, um die Diagnose zu bestätigen, so wird dies als Differential Verification Bias [68] oder Differential Reference Standard Bias [49] bezeichnet. Aufgrund der Invasivität des Referenzstandards oder des klinischen Zustands des Patienten, muss dieser Bias manchmal in Kauf genommen werden.

### **3.8.4 Diagnostic Review Bias**

Schließlich kann das Wissen um das Ergebnis des Index-Tests das Ergebnis des Referenzstandard-Tests beeinflussen [49], [68]. Dies kann zu einer Überschätzung der diagnostischen Genauigkeit eines Tests führen [49]. Daher ist es für die Bewertung einer Studie wichtig zu wissen, ob die den Test auswertenden Personen die Bewertung des Referenzstandards unabhängig vom Index-Test durchgeführt haben oder nicht.

### **3.9 Vorbereitung für die Analyse**

Aus den in die Meta-Analyse einbezogenen Studien wurden zunächst alle Werte für richtig positive, falsch positive, richtig negative und falsch negative Testergebnisse bestimmt, bzw. berechnet. Diese Werte wurden in eine Vierfeldertafel (siehe Abb. 3) übertragen, um daraus weitere interessierende Werte wie die Sensitivität und die Spezifität zu errechnen. Dies geschah wie folgt [47]:

**Richtig positive Werte (RP):** Anzahl der Patienten, bei denen der Test auf die Krankheit positiv ausfällt und die auch wirklich erkrankt ist.

**Falsch positive Werte (FP):** Anzahl der Patienten, bei denen der Test auf die Krankheit positiv ausfällt, die aber nicht erkrankt ist.

**Richtig negative Werte (RN):** Anzahl der Patienten, bei denen der Test auf die Krankheit negativ ausfällt und die auch wirklich gesund ist.

**Falsch negative Werte (FN):** Anzahl der Patienten, bei denen der Test auf die Krankheit negativ ausfällt, die aber erkrankt ist.

<b>PATIENT</b>  <b>TESTERGEBNIS</b>	Krank	Gesund
Positiv	Richtig positiv (RP)	Falsch positiv (FP)
Negativ	Falsch negativ (FN)	Richtig negativ (RN)

**Abb. 3:** Vierfeldertafel zur Bestimmung der Sensitivität und Spezifität (basierend auf Weinstein et al.

[79])

**Sensitivität:** Die Sensitivität eines Tests gibt an, mit welcher Wahrscheinlichkeit der Test bei einer erkrankten Person positiv ausfallen wird. Diese Wahrscheinlichkeit kann mit folgender Formel geschätzt werden:

$$\text{Sensitivität} = \frac{RP}{(RP+FN)} \cdot 100 \%$$

**Spezifität:** Die Spezifität eines Tests hingegen gibt an, mit welcher Wahrscheinlichkeit der Test bei einer nicht erkrankten Person negativ ausfallen wird. Eine Schätzung dieser Wahrscheinlichkeit kann berechnet werden mit:

$$\text{Spezifität} = \frac{RN}{(RN+FP)} \cdot 100 \%$$

In manchen Studien stimmten die errechneten Werte für Sensitivität und Spezifität nicht mit den Angaben der Originalpublikation überein. Gründe hierfür waren z.B. Unterschiede in der Berechnung von Mittelwerten bei mehreren Testauswertern. Für die Meta-Analyse wurde dann der neu errechnete Wert verwendet. Waren in den Studien mehrere Auswerter beteiligt und somit mehrere Werte für die Vierfeldertafel vorhanden, so wurde jeweils ein Mittelwert gebildet und dieser für die Meta-Analyse benutzt.

### **Impaktfaktoren**

Der Impaktfaktor (IF) einer Fachzeitschrift gibt an, wie häufig die Artikel der Fachzeitschrift in der wissenschaftlichen Literatur in einem bestimmten Jahr zitiert werden [66]. Er zeigt dabei auf, wie oft Artikel aus dieser Zeitschrift, die in den letzten 2 Jahren publiziert wurden, innerhalb des laufenden Jahres von anderen Zeitschriften zitiert werden und bezieht sich dabei auf alle dort veröffentlichten Artikel der letzten 2 Jahre [2], [34]:

$$IF = \frac{\text{Anzahl der Zitate im laufenden Jahr bezogen auf die Publikationen der 2 Vorjahre}}{\text{Gesamtzahl der in den letzten 2 Jahren veröffentlichten Artikel}}$$

Ursprünglich sollte der Impaktfaktor eine Hilfestellung geben, um Fachzeitschriften unabhängig von ihrer Größe miteinander vergleichen zu können [33]. Mittlerweile wird er aber auch dafür verwendet, die Autoren der Artikel miteinander zu vergleichen oder ihre Leistungen und Fähigkeiten danach zu bewerten, wie hoch der Impaktfaktor der

Zeitschriften ist, in denen ihre Artikel bisher veröffentlicht wurden [2], [32], [33]. Da der Impaktfaktor dafür aber gar nicht geeignet ist [57] und sein falscher Gebrauch schwerwiegende Folgen für die Autoren haben kann [32], [57], wird seine Verwendung kontrovers diskutiert. Auch ist der Impaktfaktor einer Zeitschrift davon abhängig, ob diese frei im Internet zugänglich ist und ihre Artikel somit leichter zitiert werden können. Artikel in einer anderen Sprache als in Englisch werden nicht mit einbezogen [2].

Zu jeder in der Meta-Analyse verwendeten Studie wurde der entsprechende Impaktfaktor der Zeitschrift des Publikationsjahres gesucht. Hierfür wurden die Webseiten der jeweiligen Fachzeitschriften oder Onlinematerial verwendet, welches basierend auf der Datenbank *Journal Citation Reports* Tabellen über Impaktfaktoren beinhaltet.

### **3.10 Bivariate logistische Regression**

Diagnostische Studien liefern als Ergebnis meist zwei Werte, die Sensitivität und die Spezifität. Diese Werte sind voneinander abhängig sind, sodass sich die Sensitivität eines diagnostischen Tests in der Regel ändern muss, wenn sich die Spezifität verändert. Andererseits wird während der gesamten bivariaten Analyse der zweidimensionale Effekt beider Werte berücksichtigt, was bei dem bisher für Meta-Analysen beliebten Receiver Operating Characteristics (ROC) nicht der Fall war [62].

Auch findet der sogenannte *Random Effect* Ansatz in der bivariaten logistischen Regression Beachtung. Dieser besagt, dass sowohl die errechnete Sensitivität als auch die Spezifität jeder Studie, die in eine Meta-Analyse eingeht, um einen gewissen Wert schwankt. Diese Schwankung ist annähernd normalverteilt [76] und kommt z.B. durch unterschiedliche Gruppengröße oder durch unerkannte Abweichungen im Studienprotokoll zustande. Zusammen mit der möglichen Korrelation zwischen Sensitivität und Spezifität bildet dieser Ansatz die Grundlage der bivariaten Analyse [62].

Um die Unterschiede der einzelnen Studien in der Analyse auszugleichen, werden die Ergebnisse unterschiedlich gewichtet. Studien mit einer großen Patientenzahl werden stärker in die Analyse einbezogen, als kleinere Studienpopulationen, da hier die



Endresultate weniger genau zu bestimmen sind. Mit der bivariaten logistischen Regression können dann die 95%-Konfidenzintervalle (KI) sowie die jeweiligen Standardabweichungen berechnet werden und die zusammengefassten logit-transformierten Sensitivitäten und Spezifitäten erfasst werden.

### 3.10.1 Formeln und Vorgehen

Für jede Studie  $I ( l, m, \dots, p )$ , die in die Meta-Analyse einbezogen wird, wird ein Wert für die Sensitivität ( $p_{A,i}$ ) und ein Wert für die Spezifität ( $p_{B,i}$ ) angegeben. Hierbei ist  $N_A$  die Anzahl der Personen mit der interessierenden Erkrankung und  $N_B$  die Anzahl der Personen ohne diese Erkrankung [62]. Die mit dem natürlichen Logarithmus gebildete logit-transformierte Sensitivität der Studie  $i$  wird mit  $\mu_{A,i}$  ( $\mu = \text{logit } p = \log \frac{p}{(1-p)}$ ) bezeichnet, die ebenso logit-transformierte Spezifität mit  $\mu_{B,i}$  [37].

Wie oben beschrieben, wird im Random Effect Modell davon ausgegangen, dass die logit-transformierten Sensitivitäten und Spezifitäten jeder Studie um einen gemeinsamen Durchschnittswert  $\mu$  normalverteilt sind. Dieser Wert wird für die Sensitivität mit  $\mu_A$  und für die Spezifität mit  $\mu_B$  angegeben, wobei  $\sigma_A^2$  die Schwankung der einzelnen Studien um den wahren Wert der Sensitivität, also ihre Varianz, bezeichnet und  $\sigma_B^2$  die Varianz der Spezifität [37], [62]. Nach *Reitsma et al. (2005)* wird in die bivariate logistische Regression explizit die mögliche Korrelation zwischen logit-transformierter Sensitivität und logit-transformierter Spezifität mit einbezogen [62]. Die Verbindung zweier Normalverteilungen, die miteinander korrelieren, führt zu folgender Formel:

$$\begin{pmatrix} \mu_{A,i} \\ \mu_{B,i} \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \Sigma \right) \text{ mit } \Sigma = \begin{pmatrix} \sigma_A^2 & \sigma_{AB} \\ \sigma_{AB} & \sigma_B^2 \end{pmatrix}$$

Mit  $\sigma_{AB}$  als Kovarianz zwischen der logit-transformierten Sensitivität und Spezifität [62].

Um nun die Unterschiede in der Erhebung von Sensitivität und Spezifität in den einzelnen Studien auszugleichen, werden diese nach der Genauigkeit bewertet, mit welcher diese Werte gemessen wurden. Wenn die Studienpopulationen und damit  $N_A$  und  $N_B$  groß sind und die zugehörige Sensitivität und Spezifität zwischen 0 und 1 liegt ( $0 < p_{A,i}; p_{B,i} < 1$ ), kann die jeweilige Varianz der zu erwartenden logit-transformierten Werte mit folgender Formel berechnet werden:

$$s_{A,i}^2 = \frac{1}{N_{A,i} \cdot p_{A,i} \cdot (1 - p_{A,i})}$$

und

$$s_{B,i}^2 = \frac{1}{N_{B,i} \cdot p_{B,i} \cdot (1 - p_{B,i})}$$

Wie bei Meta-Analysen üblich, kann man die beobachteten Varianzen der logit-transformierten Sensitivität und logit-transformierten Spezifität als festgelegte Größen ansehen und erhält somit diese Formel:

$$\begin{pmatrix} \hat{\mu}_{A,i} \\ \hat{\mu}_{B,i} \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_{A,i} \\ \mu_{B,i} \end{pmatrix}, C_i \right) \quad \text{mit } C_i = \begin{pmatrix} s_{A,i}^2 & 0 \\ 0 & s_{B,i}^2 \end{pmatrix}$$

Schlussendlich kann man für das bivariate Modell mit  $C_i$  als Diagonalmatrix, welche  $s_i^2$  beinhaltet, also schreiben:

$$\begin{pmatrix} \hat{\mu}_{A,i} \\ \hat{\mu}_{B,i} \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \Sigma + C_i \right)$$

[62].

Für die Auswertung der Analyse wurde das Statistikprogramm SAS benutzt. Als Ergebnis wurden die logit-transformierte Sensitivität ( $\mu_A$ ) und Spezifität ( $\mu_B$ ), sowie die zugehörigen 95%-Konfidenzintervalle berechnet und aus Gründen der Übersichtlichkeit und besseren Vergleichbarkeit auf eine Nachkommastelle gerundet.

Außerdem wurden verschiedenen Subgruppen in die Meta-Analyse aufgenommen. Hierzu gehörten:

- Studien der Meta-Analyse von *Colli et al. (2006)*
- Studien mit einem Evidenzgrad Level 2B
- Studien mit einem Evidenzgrad ab Level 3B
- Studien mit einem Impaktfaktor < 4
- Studien mit einem Impaktfaktor > 4
- Studien mit Lesion-by-lesion Analyse

Subgruppen, für die weniger als 3 passende Studien gefunden wurden, wurden keiner Ergebnisanalyse unterzogen.

### **3.11 Patient-by-patient- und Lesion-by-lesion-Analyse**

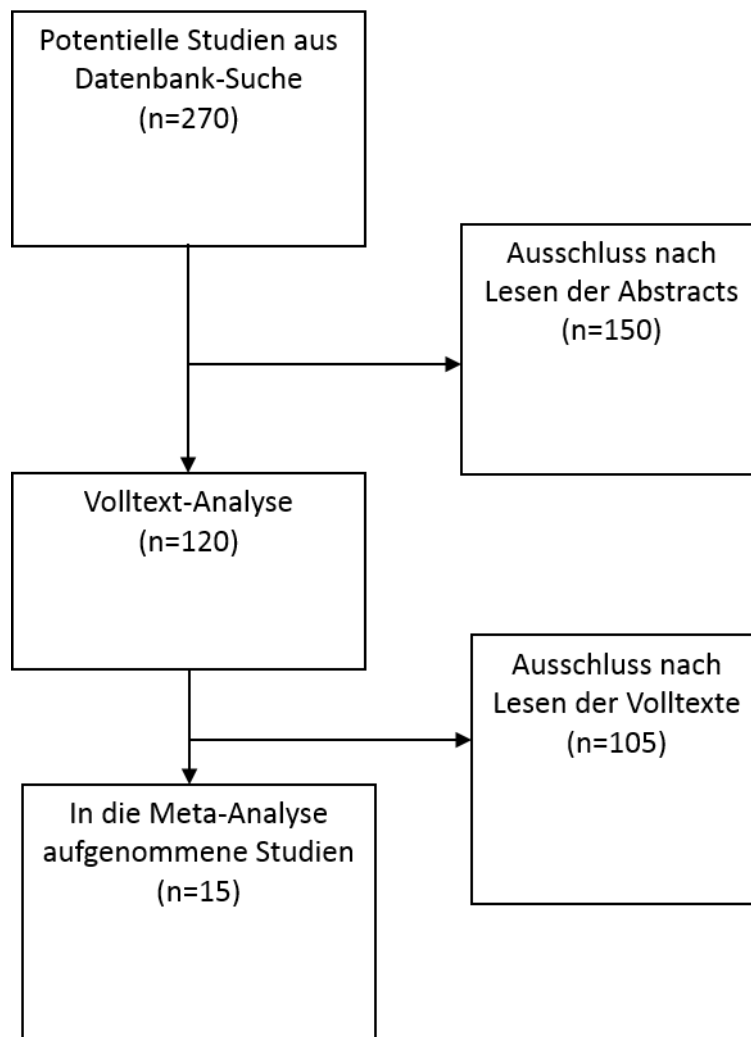
Zur Auswertung der Sensitivität und Spezifität eines bildgebenden Verfahrens in der Erkennung von Knoten in der Leber gibt es zwei unterschiedliche Herangehensweisen. Das häufiger vorkommende und für den klinischen Alltag bedeutendere Vorgehen ist die *Patient-by-patient*-Analyse. Hierbei wird als richtig positives Testergebnis angesehen, wenn bei einem an HCC erkrankten Patienten in der Bildgebung ein Tumor sichtbar ist und dieser auch im Referenzstandard als HCC bestätigt wird. Dies ist für den Patienten und für seine Weiterbehandlung von Bedeutung, weil das Ziel der Untersuchung die Unterscheidung zwischen „krank“ und „gesund“ ist. Hingegen wird bei der *Lesion-by-lesion*-Analyse darauf geachtet, wie viele Lebertumoren eines einzelnen Patienten in der Bildgebung erkannt werden. Wenn ein Patient z.B. 5 Tumoren in der Leber hat und die erste Bildgebung 3 davon erkennt, die zweite Bildgebung aber 4, dann wird der zweiten Bildgebung eine höhere Sensitivität zugeschrieben als der ersten. Da dieses Verfahren für den Patienten eine untergeordnete Bedeutung hat, weil er in beiden Fällen erkrankt ist und sich behandeln lassen muss, ist diese Analyse im klinischen Alltag selten zielführend.

Deshalb wurden diejenigen Studien, in denen die Ergebnisse nur aufgrund der Lesion-by-lesion-Analyse errechnet wurden, in der Meta-Analyse nur als Subgruppe betrachtet und separat ausgewertet. In die allgemeine Analyse für die gepoolte Sensitivität und die gepoolte Spezifität wurden diese Studien nicht aufgenommen.

## 4. ERGEBNISSE

### 4.1 Basisliteratursuche

Diese Suche ergab 270 Treffer. Von den gefundenen Studien wurden nach Lesen der Abstracts 150 Studien ausgeschlossen. Von den übrigen 120 Studien wurden die Volltexte besorgt. 15 Studien wurden in die Meta-Analyse aufgenommen (*Addley 2011, Becker-Weidman 2011, Choi SH 2008, Forner 2008, Haradome 2011, Jang 2009, Kim JE 2011, Lee MH 2011, Luca 2010, Rickes 2003, Sano 2011, Suh 2011, Xu 2008, Yoo 2009, Yu 2011*). Die Ausschlussgründe für die übrigen 105 Studien sind im Anhang 8.3 Anlage 3 als Tabelle 17 aufgeführt.

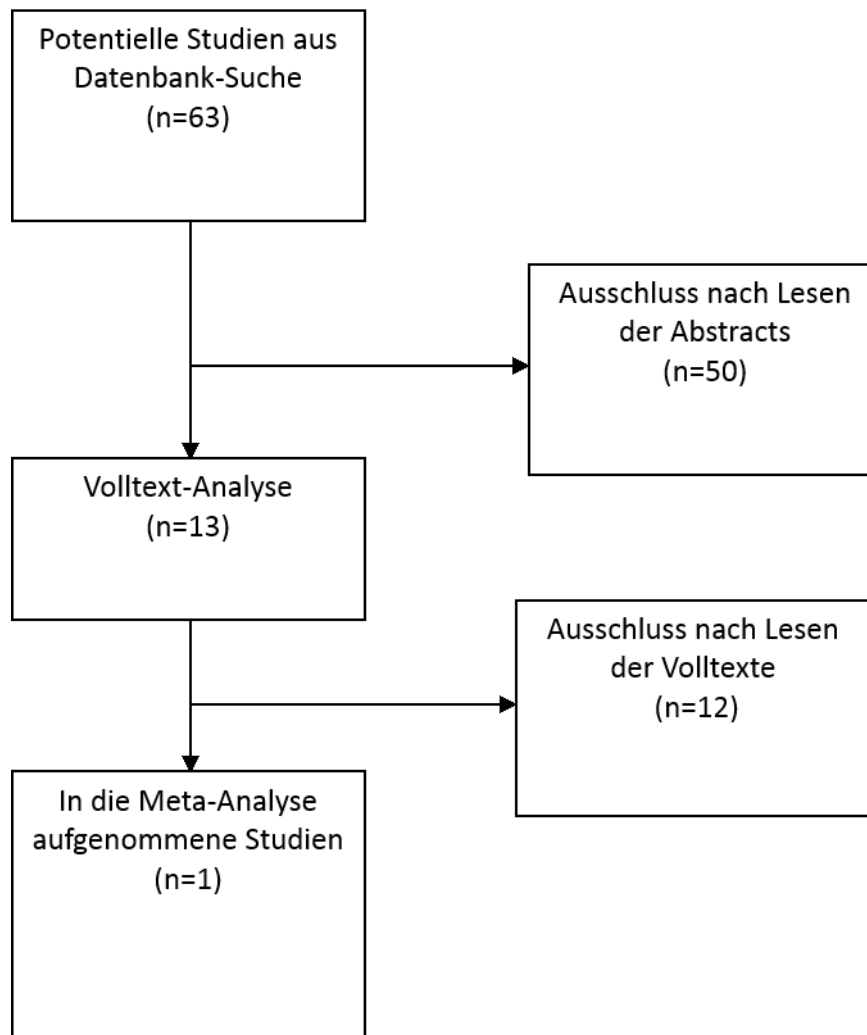


**Abb. 4:** Flow-Chart-Diagramm zum Suchvorgang der Studien der Basisliteratursuche

## 4.2 Diagnostische Studien und Ultraschall

Die Suche ergab 63 Treffer. Nach Lesen der Abstracts wurden 50 Studien ausgeschlossen. Von 13 Studien wurde der Volltext besorgt.

Nach Lesen der Studien wurden weitere 12 Studien ausgeschlossen. In 1 Studie wurden die Patienten einem sequentiellen Test unterzogen, in einer weiteren wurde keine gezielte HCC-Diagnostik betrieben. 1 Studie schloss Patienten mit Risiko für HCC (chronische Lebererkrankungen) explizit aus. 5 Studien gaben keinen geeigneten Referenzstandard an. 1 Studie gab weder Sensitivität, noch Spezifität des diagnostischen Tests an, 1 Studie führte den Referenzstandard nicht bei allen Patienten durch und 2 Studien hatten mehr als einen Ausschlussgrund. Folglich wurde 1 Studie (*Colli 2006*) aus dieser Suche in die Meta-Analyse eingeschlossen.

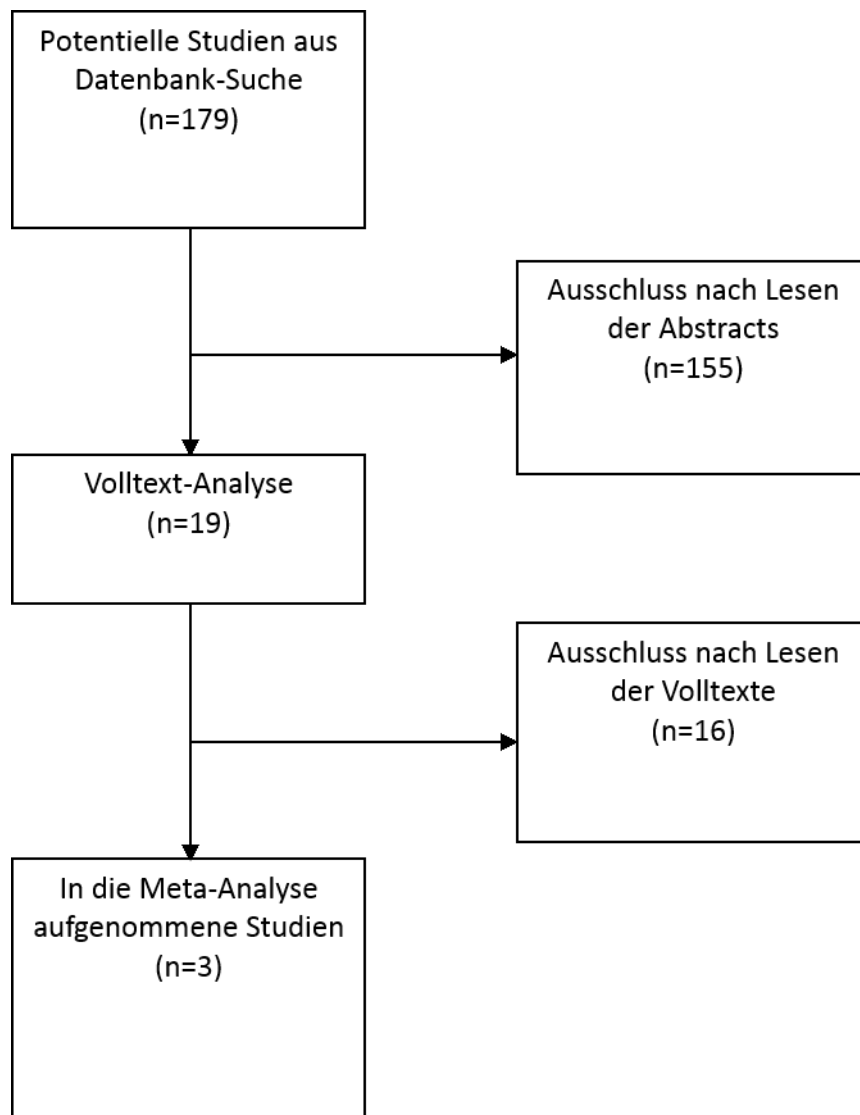


**Abb. 5:** Flow-Chart-Diagramm zum Suchvorgang der Ultraschall-Studien

### 4.3 Diagnostische Studien und CT

Diese Suchkombination ergab 179 Treffer. Das Lesen der Abstracts führte zum Ausschluss von 155 Studien. Es gab 5 Überlappungen mit dem Suchvorgang der Ultraschall-Studien. Folglich wurden von 19 Studien die Volltexte besorgt. Anschließend mussten 16 dieser Studien ausgeschlossen werden. Gründe hierfür waren die Verwendung von sequentiellen Tests (n=4), die fehlende Spezifitätsangabe (n=7), eine andere Fragestellung als die HCC-Diagnostik (n=3), der ausschließliche Einschluss von Tumoren mit einem bestimmten Durchmesser (n=1) und ein unpassender Referenzstandard (n=1).

Somit wurden 3 Studien aus diesem Suchvorgang in die Meta-Analyse eingeschlossen (*Talbot 2010, Valls 2004, Guang 2011*).

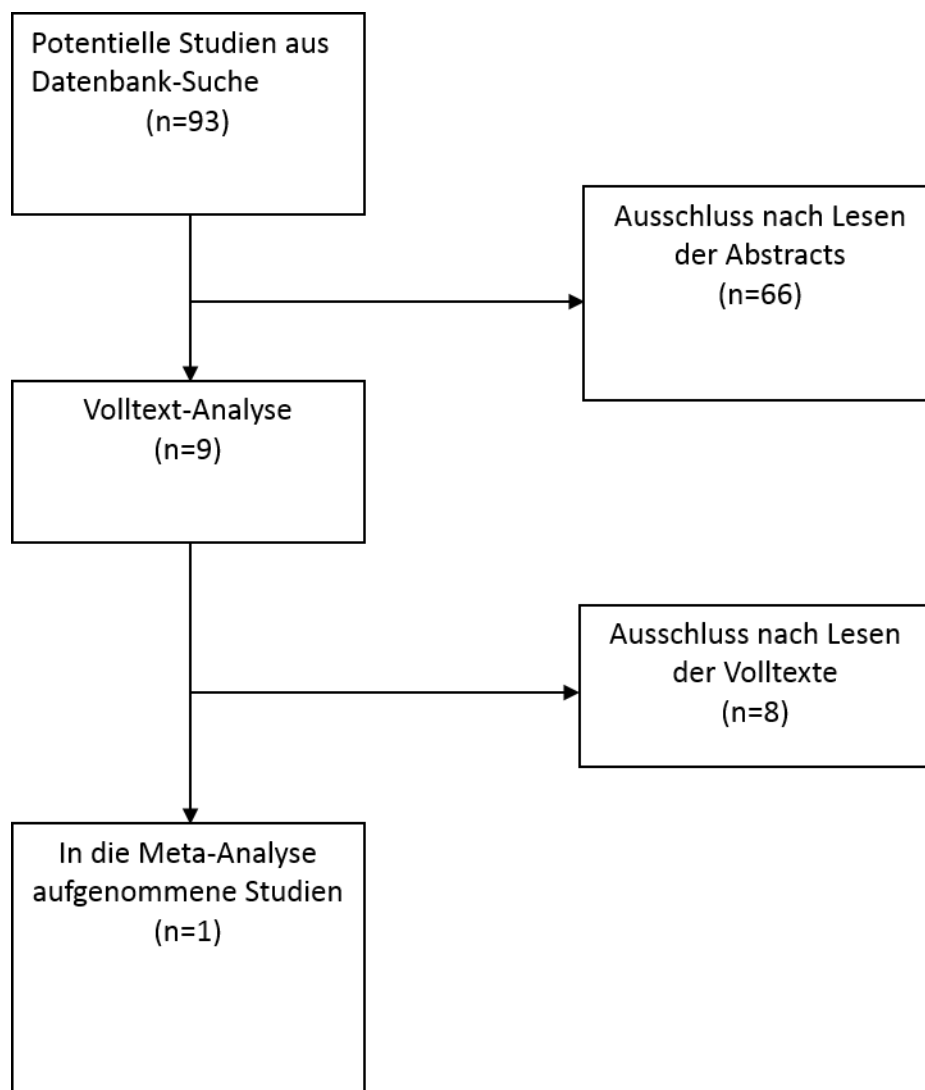


**Abb. 6:** Flow-Chart-Diagramm zum Suchvorgang der CT-Studien

#### 4.4 Diagnostische Studien und MRT

Die Suche ergab 93 Treffer. Die Abstracts von 66 Studien führten zu deren Ausschluss. Es gab 18 Überlappungen mit Studien aus den vorhergehenden Suchvorgängen. Von 9 Studien wurde der Volltext besorgt. Beim Lesen dieser Studien zeigte sich, dass 8 davon für die Meta-Analyse unbrauchbar waren. In 5 Studien lagen sequentielle Test vor. In einer Studie fehlte die Angabe der Spezifität und zwei Studien hatten mehr als einen Ausschlussgrund.

Aus der Suchkombination für MRT-Studien ging 1 Studie (*Lauenstein 2007*) in die Meta-Analyse ein.



**Abb. 7:** Flow-Chart-Diagramm zum Suchvorgang der MRT-Studien

#### **4.5 Kombinierte Suche diagnostischer Studien**

Diese Suchanfrage ergab 396 Treffer. Nach Lesen der Abstracts wurden hiervon 353 Studien ausgeschlossen. 38 Studien waren überlappend mit den vorangegangenen Suchvorgängen. Von 5 Studien wurde der Volltext besorgt. Das Lesen der Volltexte führte zum Ausschluss aller 5 Studien. 3 benutzten einen unpassenden oder unklaren Referenzstandard, 1 Studie verwendete sequentielle Tests und eine Studie gab keine Auskunft über die Spezifität des diagnostischen Tests.

Also wurde aus der kombinierten Suchanfrage keine weitere Studie in die Meta-Analyse aufgenommen.

#### **4.6 Auseinandersetzung mit den gefundenen Studien**

Insgesamt wurden, wie die Flow-Charts der Suchvorgänge zeigen, 20 Studien als für die Meta-Analyse geeignet angesehen. Bei zwei dieser Studien handelte es sich um Meta-Analysen, die ein ähnliches Ziel haben wie diese Arbeit. Der aktuellere Artikel von *Guang et al. (2011)* [36] weist einige Unterschiede auf, sodass die hierin enthaltenen Studien nicht in die Meta-Analyse übernommen wurden. Gründe hierfür waren einerseits die Fokussierung auf das Kontrastmittel SonoVue® für die Ultraschalluntersuchung, andererseits wurde in den ausgewählten Studien nicht nur das HCC diagnostiziert sondern alle fokalen Leberläsionen, sodass eine grundsätzlich andere Studienauswahl getroffen wurde.

Die ältere Meta-Analyse von *Colli et al. (2006)* [M10] erwies sich als gut zum Thema passend und sorgfältig durchgeführt. Da die Ergebnisse dieser Studie schon einige Jahre zurückliegen und die methodische Analyse damals eine andere war als die bivariate logistische Regression aus dieser Meta-Analyse, wurden die Daten und Studien von *Colli et al. (2006)* zusätzlich aufgenommen und erneut ausgewertet um zu testen, ob sich das gleiche Ergebnis zeigt und um herauszufinden, ob sich deren Studienergebnisse von denen in dieser Meta-Analyse unterscheiden.

Somit wurde von ursprünglich 20 Artikeln einer aussortiert [36] und aus einer Meta-Analyse [M10] wurden 28 Studien übernommen, ohne diese Arbeit selbst in die



Auswertung mit aufzunehmen, sodass insgesamt 46 Artikel die Grundlage dieser Meta-Analyse bilden.

#### **4.7 Eigenschaften der aufgenommenen Studien**

Von insgesamt 46 aufgenommenen Arbeiten gaben 9 Studien ihre Ergebnisse nur mit der Lesion-by-Lesion Auswertung an, davon waren 2 Studien aus der Meta-Analyse von 2006. Da die Lesion-by-Lesion Angabe für den Patienten unerheblich ist, wurde ihr eine geringere Rolle zugeschrieben. Deshalb wurden diese Studien im ersten Durchgang der Berechnung nicht beachtet. Sie wurden einer eigenen Analyse unterzogen.

Die restlichen Studien wurden in der Analyse zusammengefasst, sodass insgesamt ein Patientenkollektiv von 10924 Personen zustande kam.

##### **4.7.1 Studien mit Ultraschall**

Insgesamt wurden 19 Studien gefunden, die den Ultraschall für die HCC-Diagnostik verwendeten. Davon waren 14 Studien prospektiv ([M12], [M13], [M16], [M19], [M23], [M27], [M32], [M35], [M36], [M37], [M41], [M42], [M43], [M46]) und 5 Studien waren retrospektiv ([M4], [M8], [M14], [M17], [M48]). Insgesamt wurden in den Studien, die ihre Ergebnisse in der Patient-by-patient-Analyse angaben, 7927 Patienten mit Ultraschall untersucht. Die Anzahl der Patienten reichte von 25 [M42] bis 5339 [M41]. Die Hälfte der Studien untersuchte Patienten mit chronischer Lebererkrankung, v.a. Leberzirrhose, vor einer geplanten Lebertransplantation ([M4], [M8], [M12], [M14], [M17], [M23], [M37], [M42], [M48]). In 4 Studien wurde der Ultraschall bei Risikopatienten für HCC durchgeführt ([M13], [M16], [M19], [M43]), bei 3 Studien wurde bei Patienten aufgrund erhöhter Laborwerte wie z.B. AFP, typischer Symptome oder einer verdächtigen Bildgebung der Verdacht auf eine Neoplasie geäußert und Ultraschalluntersuchung durchgeführt ([M27], [M32], [M35]). In einer Studie wurden Patienten mit fokaler Leberläsion aufgenommen [M46] und eine Studie schloss alle Patienten ein, die im Jahr 1980 einen Ultraschall des Oberbauchs bekommen hatten und in Osaka lebten [M41].

In der Studie von *Bennett et al.* [M4] wurden 200 Patienten untersucht, jedoch wurden die Angaben über die diagnostische Richtigkeit nur auf 57 Patienten mit im Ultraschall

sichtbaren Läsionen angewendet. Hierbei waren 27 Patienten an HCC erkrankt und 30 hatten andere Leberläsionen. Als Gesamtkollektiv wurden deshalb 57 Patienten angesehen.

Der Großteil der Studien führte einen Ultraschall des Bauches ohne die Unterstützung von Doppler-Einstellungen oder Kontrastmittel durch. Zwei Studien verwendeten den dopplergestützten US ([M23], [M35]), zwei verwendeten beide Techniken ([M14], [M42]). Nur in drei Studien wurde der Kontrastmittel-verstärkte US benutzt ([M13], [M16], [M46]), in zwei Studien SonoVue®([M13], [M46]) und in einer Studie Definity® [M16].

Bei der Studie von *Teefey et al.* [M42] wurde die Auswertung des diagnostischen Tests von zwei Untersuchern vorgenommen. Für die Ultraschalluntersuchung fehlten bei einem der Untersucher die Testergebnisse von zwei nicht-HCC-Patienten. Da diese notwendig für die Berechnung der Spezifität sind, wurden hier nur die Werte des ersten Untersuchers in die Analyse aufgenommen.

6 der Ultraschall-Studien wurden mit einem Evidenzgrad 2B bewertet ([M12], [M13], [M35], [M43], [M46], [M48]). Dieser bildete den höchsten Evidenzgrad in dieser Gruppe. 7 Studien wiesen einen Evidenzgrad 3B auf ([M4], [M8], [M16], [M23], [M27], [M37], [M42]) und 4 Studien bekamen den Evidenzgrad 4 ([M14], [M17], [M19], [M32]) zugeordnet. Die Lesion-by-lesion-Analyse wurde in 1 Studie mit 43 Patienten und 69 Knoten angewendet [M36]. Ihre Ergebnisse wurden nicht in der Gesamtauswertung der Meta-Analyse aufgenommen. Da eine einzelne Studie keine Subgruppe bilden kann, wurde auch auf diese Auswertung verzichtet.

#### **4.7.2 Studien mit Computertomographie**

18 Studien untersuchten die CT auf ihre diagnostische Aussagekraft zur Diagnostik des HCC. Darunter fanden sich 9 prospektive ([M11], [M23], [M24], [M33], [M36], [M40], [M42], [M45], [M49]) und 9 retrospektive Studien ([M1], [M6], [M8], [M14], [M15], [M26], [M30], [M38], [M48]). 13 Studien stellen ihre Ergebnisse mit der Patient-by-patient-Analyse dar. Sie geben Aufschluss über die Untersuchungen von 1578 Patienten. Die Stichprobenzahl der einzelnen Studien reichte von 16 [M23] bis 430 [M33] Patienten.

In 13 Studien wurden Patienten mit chronischer Lebererkrankung vor einer Lebertransplantation mit CT-Bildgebung untersucht ([M1], [M6], [M8], [M11], [M14], [M23], [M24], [M30], [M33], [M40], [M42], [M45], [M48]). Nur in einer Studie wurden Patienten mit chronischer Lebererkrankung mit CT untersucht, bei denen Knoten in der Leber vermutet wurden [M40].

Die angewendeten CT-Techniken der Studien waren sehr unterschiedlich. In den meisten Fällen wurde aber ein Kontrastmittel angewendet, hauptsächlich für die dual- oder triphasische Bildgebung ([M6], [M8], [M11], [M23], [M24], [M30], [M42], [M45]). Lediglich eine Studie verwendete nur das konventionelle CT [M14] und eine Studie wandte das PET-CT an [M40]. In 2 Studien kam das Multidetector-CT zum Einsatz ([M1], [M48]), in den anderen v.a. das Spiral-CT.

In der Studie von *Burrel et al. (2003)* [M7] wurde nicht bei allen Patienten eine CT-Untersuchung durchgeführt. Deshalb wurde diese Studie nur in die MRT-Auswertung mit einbezogen, weil alle Patienten eine MRT-Untersuchung bekamen. Die Daten für die CT-Auswertung wurden nicht übernommen.

In der Studie von *Teefey et al. (2003)* [M42] wurde die diagnostische Auswertung von zwei Untersuchern durchgeführt, sodass jeweils zwei Ergebnisse vorlagen. Für die Sensitivität und die Spezifität wurde jeweils der Mittelwert errechnet und das Ergebnis ggf. aufgerundet. Durch Abrunden des korrelierenden falsch negativen bzw. falsch positiven Wertes wurde die Gesamtpatientenzahl berücksichtigt.

In 12 von 13 Studien handelte es sich ausschließlich um HCC-Patienten. Bei einer Studie wurden maligne Leberläsionen und nicht das hepatozelluläre Karzinom als interessierende Erkrankung gesucht [M30]. Da es sich hierbei um eine Studie aus der Meta-Analyse *Colli et al. (2006)* handelt, wurde sie aus Gründen der Vollständigkeit trotzdem in die Berechnung dieser Meta-Analyse aufgenommen.

Von 13 analysierten Studien hatten 5 mit einem Evidenzgrad 2B den höchsten aller CT-Studien ([M1], [M11], [M24], [M33], [M48]). Der Evidenzgrad 3B war in dieser Gruppe mit 7 Studien am häufigsten vertreten ([M6], [M8], [M23], [M30], [M40], [M42], [M45]). Eine Studie erhielt den Evidenzgrad 4 [M14].

1 Studie gab die Ergebnisse sowohl in Patient-by-patient-Analyse als auch in der Lesion-by-lesion-Analyse an [M45]. In diesem Fall wurden nur erstere Daten verwendet und in die Gesamtanalyse aufgenommen. 5 Studien gaben ihre Ergebnisse ausschließlich in Form der Lesion-by-lesion-Analyse an ([M15], [M26], [M36], [M38], [M49]). Diese Studien umfassten 330 Patienten mit 651 Leberläsionen. Die Studien wurden als Subgruppe analysiert, jedoch nicht in die Gesamtanalyse einbezogen.

#### **4.7.3 Studien mit Magnetresonanztomographie**

In 20 Studien wurde von den Patienten eine MRT-Bildgebung gemacht. Davon haben 10 Studien ein prospektives Design ([M5], [M7], [M11], [M13], [M20], [M21], [M23], [M29], [M36], [M42]) und 10 ein retrospektives ([M3], [M6], [M9], [M15], [M18], [M22], [M38], [M39], [M47], [M48]) verwendet. 13 Studien gaben ihre Ergebnisse in Form der Patient-by-patient-Analyse an. Insgesamt wurden hierbei 882 Patienten mit MRT untersucht. Die Anzahl der Patienten reichte von 25 [M42] bis 216 [M48].

In allen 13 Studien wurden MRT-Aufnahmen von Patienten gemacht, die sich aufgrund einer chronischen Lebererkrankung einer Organtransplantation unterzogen.

Fast alle aufgenommenen Studien benutzten ein Kontrastmittel, das Gadolinium oder ein Gadoliniumchelat beinhaltet. So z.B. Gadopentetat-Dimeglumin ([M6], [M11], [M20], [M23]), Gadodiamid ([M13], [M42], [M48]), Gadobenat-Dimeglumin ([M3], [M9]) oder Gadolinium-Chelat [M21]. In einer Studie wurden sogar 2 verschiedene Kontrastmittel angewendet, zum einen superparamagnetische Eisenoxid-Partikel und zum anderen entweder Gadopentetat-Dimeglumin oder Gadodiamid [M5]. In der Studie von *Libbrecht et al. (2002)* ([M23]) wurde entweder Gadopentetat-Dimeglumin oder Meglumingadoterat verwendet. Nur die Studie von *Mori et al. (2002)* verwendete kein Gadolinium-haltiges Kontrastmittel, sondern ausschließlich superparamagnetische Eisenoxide [M29].

In allen Studien wurden mindestens die Standard T1- und T2-Gewichtung für die MRT-Bildgebung gewählt. Meistens wurden die Bilder vor und unmittelbar nach Kontrastmittelgabe gewonnen. Zur genaueren Darstellung der Leber wurden die Patienten häufig gebeten, während der Aufnahmen die Luft anzuhalten (sog. *Breath hold* – Technik).

In 2 von 13 Studien wurden nicht nur HCC-Patienten oder Patienten mit Verdacht auf HCC im MRT untersucht, sondern auch Patienten mit anderen Malignomen der Leber [M20], [M29]. Auch hierbei handelt es sich um Studien aus der Meta-Analyse von *Colli et al. (2006)*, sodass auch diese beiden Studien in der Gesamtanalyse enthalten sind.

In der Studie von *Teefey et al. [M42]* wurden die Ergebnisse des diagnostischen Tests wieder von zwei Untersuchern durchgeführt. Da die für die Berechnung der Sensitivität notwendigen Werte (richtig positive Werte und falsch negative Werte) nur von einem der Untersucher vollständig vorlagen, wurden auch nur diese für die Berechnung der Sensitivität in die Analyse mit aufgenommen.

Keine der Studien wies einen Evidenzgrad höher als 2B auf. Diesen hatten 6 Studien ([M7], [M11], [M13], [M20], [M21], [M48]). Die restlichen 7 Studien hatten den Evidenzgrad 3B.

7 Studien mit 498 Patienten und 831 Leberläsionen gaben ihre Ergebnisse ausschließlich in der Lesion-by-lesion-Analyse wieder ([M15], [M18], [M22], [M36], [M38], [M39], [M47]). Auch hier wurden die Ergebnisse wieder als Subgruppe analysiert, jedoch nicht in der Gesamtanalyse ausgewertet.

#### **4.7.4 Studien mit AFP**

Die Studien zur AFP-Diagnostik wurden alle aus der Meta-Analyse von *Colli et al (2006)* [M10] übernommen, da keine gesonderte Literatursuche zum Thema „diagnostische Studien und AFP“ durchgeführt wurde (siehe 3.6 Diagnostische Studien und AFP).

Folglich wurden 10 Studien gefunden, die AFP in der Diagnostik des HCC untersuchen. Davon waren 3 prospektiv ([M32], [M34], [M43]), 2 retrospektiv ([M8], [M14]) und 5 waren Fall-Kontroll-Studien ([M2], [M25], [M28], [M31], [M44]). Insgesamt wurden die AFP-Werte von 3453 Patienten gemessen. Die Stichprobenzahl der einzelnen Studien reichte von 106 [M14] bis 1099 [M34]. 2 der Studien konnten aufgrund fehlender Vergleichbarkeit nicht in die Analyse mit aufgenommen werden (s.u.) [M2], [M34]. Somit gingen 8 Studien mit insgesamt 2154 Patienten in die Berechnung der AFP-Diagnostik mit ein.

Die Patientenkollektive der einzelnen Studien waren sehr unterschiedlich. 2 Studien untersuchten Patienten mit chronischer Lebererkrankung vor der Transplantation ([M8], [M14]), 1 Studie untersuchte Risikopatienten für die Entwicklung von HCC [M43], 1 Studie untersuchte Patienten mit chronischer Lebererkrankung und/oder einem erhöhten AFP-Wert in der Routineuntersuchung [M32], 2 Studien formten Fall-Kontroll-Gruppen aus Patienten mit HCC und einer chronischen Lebererkrankung als Fällen und Patienten mit chronischer Lebererkrankung als Kontrollen ([M31], [M44]) und die restlichen 4 Studien formten andere Fall-Kontroll-Gruppen mit 2-4 Untergruppen ([M2], [M25], [M28], [M34]).

Da die Cut-off-Werte, die zur Unterscheidung zwischen erhöhtem AFP und AFP-Normalwert dienten, von Studie zu Studie stark variierten, wurden in die Auswertung nur Studien eingeschlossen, die einen möglichst ähnlichen Cut-off-Wert benutzten. Somit wurden in einer ersten Analyse 7 Studien mit einem Cut-off-Wert von AFP>20 ng/ml eingeschlossen ([M8], [M14], [M28], [M31], [M32], [M43], [M44]) und in einer zweiten Analyse 5 Studien, bei denen ein AFP-Wert >200 ng/ml als verdächtig galt ([M8], [M25], [M31], [M32], [M44]). Die Studien von *Piantino et al.* [M34] und von *Bayati et al.* [M2] verwendeten andere Cut-off-Werte, sodass deren Ergebnisse nicht in die Meta-Analyse mit aufgenommen wurden.

Nur einer AFP-Studie wurde der Evidenzgrad 2B zugeteilt [M43]. Der Evidenzgrad 3B wurde 6-mal verteilt ([M8], [M25], [M28], [M31], [M34], [M44]) und der Evidenzgrad 4 wurde an 3 Studien vergeben ([M14], [M32], [M49]).

## 4.8 Daten der aufgenommenen Studien

**4.8.1 Tabelle 2: Daten der Studien mit US als diagnostischem Test**

Autor/Jahr	Land	Patientenkollektiv	US-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Forner, 2008 [M13]	Spanien	Asymptomatische Patienten mit Zirrhose, keine HCC-Vorgeschichte, 1 solitärer Knoten neu entdeckt im US	CEUS	89	31	2	29	27	52	93	Histopathologie aus Biopsat
Jang, 2009 [M16]	Kanada	Patienten mit Hepatitis, Leberknoten 1-2 cm Ø	CEUS	59	26	0	4	29	87	100	Histologie oder Leberexplantat; Follow-up mind. 12 Monate
Rickes, 2003 [M35]	Deutschland	Patienten mit Zirrhose, Tumorverdacht durch Symptomatik, Bildgebung oder Labor	Doppler	87	58	9	7	13	89	54	Histologie; AFP≥500 ng/ml und Follow-up mind. 18 Monate

Autor/Jahr	Land	Patientenkollektiv	US-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Xu, 2007 [M46]	China	Patienten ohne Vorbehandlung, keine Zysten, Tumor $\leq 2\text{cm } \emptyset$	CEUS	104	39	4	10	51	80	93	Histologie; Bildgebung laut EASL-Kriterien
Yu, 2011 [M48]	USA	Patienten mit chronischer Lebererkrankung vor Transplantation	Nicht kontrastmittel-verstärkt (unenanced)	430	88	11	50	281	64	96	Leberexplantat
Okazaki, 1984 [M32]	Japan	Asymptomatische Patienten mit chronischer Lebererkrankung, AFP $\uparrow$	Nicht kontrastmittel-verstärkt (unenanced)	245	12	3	2	228	86	99	Histologie; klinische Diagnostik
Tong, 2001 [M43]	USA	Patienten mit Hepatitis B, C oder beide, an Klink überwiesen	Nicht kontrastmittel-verstärkt (unenanced)	602	31	11	0	560	100	98	Histologie, Follow-up mind. 12 Monate
Chalasani, 1999 [M8]	USA	Patienten mit Zirrhose vor Transplantation	Nicht kontrastmittel-verstärkt (unenanced) + Doppler	285	16	18	11	240	59	93	Histologie; in einem Fall: typische Bildgebung



<b>Autor/Jahr</b>	<b>Land</b>	<b>Patientenkollektiv</b>	<b>US-Untersuchung</b>	<b>N</b>	<b>RP</b>	<b>FP</b>	<b>FN</b>	<b>RN</b>	<b>SENS</b>	<b>SPEZ</b>	<b>Referenzstandard</b>
Gambarin-Gelwan, 2000 [M14]	USA	Patienten mit Zirrhose vor Transplantation	Nicht kontrastmittel-verstärkt (unenanced)	106	11	5	8	82	58	94	Leberexplantat
Kobayashi, 1985 [M19]	Japan	Patienten mit Zirrhose, Screening	Nicht kontrastmittel-verstärkt (unenanced)	95	6	2	2	85	75	98	Histologie, Follow-up bis 12 Monate
Dodd, 1992 [M12]	USA	Patienten mit Zirrhose vor Transplantation	Nicht kontrastmittel-verstärkt (unenanced)	200	12	3	16	169	43	98	Leberexplantat
Saada, 1997 [M37]	England	Patienten mit chronischer Lebererkrankung, vor Transplantation	Nicht kontrastmittel-verstärkt (unenanced)	39	2	0	4	33	33	100	Leberexplantat
Kim CK, 2001 [M17]	Südkorea	Patienten mit Zirrhose vor Transplantation	Nicht kontrastmittel-verstärkt (unenanced)	52	6	3	10	33	38	92	Leberexplantat

Autor/Jahr	Land	Patientenkollektiv	US-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Bennett, 2002 [M4]	USA	Patienten mit Zirrhose vor Transplantation	Nicht kontrastmittel- verstärkt (unenhanced)	57	8	5	19	16	30	76	Leberexplantat
Teefey, 2003 [M42]	USA	Patienten mit chronischer Lebererkrankung vor Transplantation + AFP ↑ oder primär sklerosierende Cholangitis	Nicht kontrastmittel- verstärkt (unenhanced) + Doppler	25	8	4	1	12	89	75	Leberexplantat
Libbrecht, 2003 [M23]	Belgien	Patienten mit Zirrhose vor Transplantation	Farbdoppler	46	6	0	9	31	40	100	Leberexplantat
Maringhini, 1984 [M27]	Italien	Patienten mit Zirrhose Verdacht auf maligne Neoplasie (Symptomatik, Labor)	Nicht kontrastmittel- verstärkt (unenhanced)	67	22	6	2	37	92	86	Histologie aus Biopsat

Autor/Jahr	Land	Patientenkollektiv	US-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Tanaka, 1986 [M41]	Japan	Patienten aus dem Amtsbezirk Osaka, die US des oberen Bauchraums bekamen	Nicht kontrastmittel- verstärkt (unenhanced)	5339	67	5	46	5221	59	99,9	Follow-up bis 36 Monate; Histologie
Rode, 2001 [M36]	Frank- reich	Patienten mit Zirrhose, vor Transplantation, ohne Vorbehandlung	Farbdoppler	43P/ 69K *	6 *	53 *	3 *	7 *	46 *	95 *	Leberexplantat

\* = Lesion-by-lesion-Angaben, P= Patienten, K= Knoten

#### 4.8.2 Tabelle 3: Daten der Studien mit CT als diagnostischem Test

Autor /Jahr	Land	Patientenkollektiv	CT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Addley, 2011 [M1]	England	Patienten vor Leber- transplantation, keine Vorbehandlung	MDCT	39	21	3	8	7	72	70	Leberexplantat
Yu, 2011 [M48]	USA	Patienten mit chronischer Lebererkrankung vor Transplantation	SDCT/MDCT	385	113	9	36	227	76	96	Leberexplantat

Autor /Jahr	Land	Patientenkollektiv	CT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Valls, 2004 [M45]	Spanien	Patienten mit Zirrhose, vor Transplantation, keine Vorbehandlung	Dual-phase helical CT	85	39	7	12	27	76	79	Leberexplantat
Talbot, 2010 [M40]	Frank- reich	Patienten mit chronischer Lebererkrankung, ohne Vorbehandlung	F-FDG- u. F- Fluorocholine- PET/CT	51	23	1	11	16	68	94	Histologie; EASL Kriterien
Teefey, 2003 [M42]	USA	Patienten mit chronischer Lebererkrankung vor Transplantation + AFP ↑ oder primär sklerosierende Cholangitis	Kontrastmittel- verstärktes CT (Optiray®)	25	6	4	3	12	67	75	Leberexplantat
Chalasani, 1999 [M8]	USA	Patienten mit Zirrhose vor Transplantation	Dual-phase Spiral- CT	285	25	10	2	248	93	96	Histologie; in einem Fall: typische Bildgebung

<b>Autor /Jahr</b>	<b>Land</b>	<b>Patientenkollektiv</b>	<b>CT-Untersuchung</b>	<b>N</b>	<b>RP</b>	<b>FP</b>	<b>FN</b>	<b>RN</b>	<b>SENS</b>	<b>SPEZ</b>	<b>Referenzstandard</b>
Gambarin-Gelwan, 2000 [M14]	USA	Patienten mit Zirrhose vor Transplantation	Konventionelles CT	106	10	5	9	82	53	94	Leberexplantat
Libbrecht, 2003 [M23]	Belgien	Patienten mit Zirrhose vor Transplantation	Kontrastmittel- verstärktes helical CT (Omnipaque® oder Iomeron®)	16	1	3	1	11	50	79	Leberexplantat
Lim, 2000 [M24]	Südkorea	Patienten mit chronischer Lebererkrankung, vor Transplantation	Triphasisches helical CT	41	12	1	3	25	80	96	Leberexplantat
Peterson, 2000 [M33]	USA	Asymptomatische Patienten mit chronischer Lebererkrankung, vor Transplantation	Helical CT	430	26	30	33	341	44	92	Leberexplantat
Mortelé, 2001 [M30]	Belgien	Patienten mit chronischer Lebererkrankung, vor Transplantation	Dual-phase Spiral- CT	53	14	0	3	36	82	100	Leberexplantat

Autor /Jahr	Land	Patientenkollektiv	CT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
De Lédinghen, 2002 [M11]	Frank- reich	Patienten mit Zirrhose, vor Transplantation	Dual-phase helical CT	34	17	2	4	11	81	85	Leberexplantat
Born, 1998 [M6]	Deutsch- land	Patienten mit Zirrhose, vor Transplantation	Spiral-CT, Kontrastmittel- verstärktes CTAP (Ultravist®)	28	10	1	4	13	71	93	Histologie
Zacherl, 2002 [M49]	Öster- reich	Patienten mit chronischer Lebererkrankung, vor Transplantation	Multiphasisches helical CT	23P/ 73K *	37 *	13*	10*	13 *	74 *	57 *	Leberexplantat
Rode, 2001 [M36]	Frank- reich	Patienten mit Zirrhose, vor Transplantation, ohne Vorbehandlung	Helical CT	43P/ 69K *	7 *	52*	4 *	6 *	54 *	93 *	Leberexplantat
Haradome, 2011 [M15]	Japan/ Italien	Patienten mit Verdacht auf HCC und CT sowie MRT Untersuchung	MDCT	75P/ 99K *	42 *	37*	2 *	18 *	70 *	95 *	Histologie aus Resektat bzw. Biopsat

Autor /Jahr	Land	Patientenkollektiv	CT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Luca, 2010 [M26]	Italien	Patienten mit Zirrhose, vor Transplantation	MDCT	125P / 158K *	116 *	21 *	6 *	15 *	89 *	78 *	Leberexplantat
Sano, 1997 [M37]	Japan	Patienten nach Leberresektion	Kontrastmittel- verstärktes CT (Omnipaque®, Iomeron®, Iopamiron®)	64P/ 252K *	58 *	158 *	3 *	33 *	64 *	98 *	Histologie aus Resektat

SDCT= Single detector CT, MDCT= multi detector CT;

\* = Lesion-by-lesion Angaben, P=Patienten, K=Knoten

#### 4.8.3 Tabelle 4: Daten der Studien mit MRT als diagnostischem Test

Autor/Jahr	Land	Patientenkollektiv	MRT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Becker- Weidmann, 2011 [M3]	USA	Patienten mit chronischer Lebererkrankung, vor Transplantation	Gadobenate- dimeglumine MRT	101	34	0	1	66	97	100	Leberexplantat
Choi, 2008 [M9]	Süd-korea	Patienten mit Zirrhose, vor Transplantation	Gadobenate- dimeglumine MRT	47	19	5	4	19	83	79	Leberexplantat

Autor/Jahr	Land	Patientenkollektiv	MRT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Forner, 2008 [M13]	Spanien	Asymptomatische Patienten mit Zirrrose, keine HCC-Vorgeschichte, 1 solitärer Knoten neu entdeckt im US	Gadodiamide-MRT	89	37	1	23	28	62	97	Histopathologie aus Biopsat
Yu, 2011 [M48]	USA	Patienten mit chronischer Lebererkrankung vor Transplantation	Gadodiamide-MRT	216	99	13	18	86	85	87	Leberexplantat
Lauenstein, 2007 [M21]	USA	Patienten mit chronischer Lebererkrankung vor Transplantation	Gadolinium-chelate MRT	115	24	2	3	86	89	98	Leberexplantat
Teefey, 2003 [M42]	USA	Patienten mit chronischer Lebererkrankung vor Transplantation + AFP ↑ oder primär sklerosierende Cholangitis	Gadodiamide-MRT	25	5	4	4	12	56	75	Leberexplantat



<b>Autor/Jahr</b>	<b>Land</b>	<b>Patientenkollektiv</b>	<b>MRT-Untersuchung</b>	<b>N</b>	<b>RP</b>	<b>FP</b>	<b>FN</b>	<b>RN</b>	<b>SENS</b>	<b>SPEZ</b>	<b>Referenzstandard</b>
Libbrecht, 2003 [M23]	Belgien	Patienten mit Zirrhose vor Transplantation	Dimeglumine- gadopentetate/Meglu- mine-gadoterate-MRT	27	7	3	3	14	70	82	Leberexplantat
De Lédinghen, 2002 [M11]	Frank- reich	Patienten mit Zirrhose, vor Transplantation	Gadopentetate- dimeglumine-MRT	34	19	0	2	13	90	100	Leberexplantat
Born, 1998 [M6]	Deutsch- land	Patienten mit Zirrhose, vor Transplantation	Gadolinium-DTPA- MRT	28	11	1	5	11	69	92	Histologie
Mori, 2002 [M29]	Deutsch- land	Patienten mit chronischer Lebererkrankung vor Transplantation	Ferumoxide + Glucose	48	17	7	3	21	85	75	Leberexplantat
Bhartia, 2002 [M5]	England	Patienten mit Zirrhose, vor Transplantation	Gadopentetate- dimeglumine/Gadodia mide MRT	31	13	0	1	17	93	100	Leberexplantat
Burrel, 2003 [M7]	Spanien	Patienten mit Zirrhose, vor Transplantation	Gadolinium-MRT	50	29	1	0	20	100	95	Leberexplantat
Krinsky, 2001 [M20]	USA	Patienten mit Zirrhose, vor Transplantation	Gadopentetate- dimeglumine MRT	71	6	10	5	50	55	83	Leberexplantat

Autor/Jahr	Land	Patientenkollektiv	MRT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Rode, 2001 [M36]	Frank- reich	Patienten mit Zirrhose, vor Transplantation, ohne Vorbehandlung	Gadolinium-MRT	43P/ 69K *	10 *	32 *	24*	3 *	77 *	57 *	Leberexplantat
Suh, 2011 [M39]	Südkorea	Patienten mit bekannten oder vermuteten fokalen Leberläsionen nach Gadoxet-verstärkter MRT-Bildgebung mit hyperintensen Knoten in der hepatobiliären Phase	Gadoxetsäure-MRT	48P/ 48K *	9 *	32 *	0 *	7 *	56 *	100 *	Histologie, Follow-up mind. 6 Monate, EASL Kriterien in MRT-Bildgebung
Haradome, 2011 [M15]	Japan/ Italien	Patienten mit Verdacht auf HCC und statt gehabter CT sowie MRT Untersuchung	Gadolinium-EOB-DTPA verstärktes MRT	75P/ 99K *	46 *	36 *	3 *	14*	77 *	92 *	Histologie aus Resektat bzw. Biopsat
Sano, 2011 [M38]	Japan	Patienten nach Leberresektion	Gadoxetsäure-MRT	64P/ 252K *	87 *	154 *	7 *	4 *	96 *	96 *	Histologie aus Resektat

Autor/Jahr	Land	Patientenkollektiv	MRT-Untersuchung	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Lee, 2011 [M22]	Südkorea	Patienten mit histologisch bestätigtem HCC und benignen Leberläsionen, Gadolinium-verstärktes MRT	Gadoxetsäure-MRT	66P/ 70K *	39 *	10 *	14 *	7 *	85 *	42 *	Histologie
Kim JE, 2011 [M18]	Südkorea	Patienten mit Risiko für oder Zustand nach HCC, Gadolinium-verstärktes MRT	Gadoxetsäure-MRT	94P/ 141K *	73 *	29 *	4 *	35 *	68 *	88 *	Histopathologie, Follow-up mind. 15 Monate
Yoo, 2009 [M47]	Südkorea	Patienten mit Zirrhose, vor Resektion oder Transplantation	Superpara-magnetische Eisenoxide (SPIO) MRT	108P / 152K *	108 *	25 *	3 *	16 *	87 *	89 *	Histopathologie aus Resektat bzw. Explantat

\* = Lesion-by-lesion Angaben, P = Patienten, K = Knoten

**4.8.4 Tabelle 5: Daten der Studien mit AFP als diagnostischem Test**

Autor/Jahr	Land	Patientenkollektiv	AFP-Cut-off-Wert	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Trevisani, 2001 [M44]	Italien	170 HCC-Patienten, 170 Kontrollen mit chron. Lebererkrankung	>20 ng/ml	340	102	16	68	154	60	90,6	Histologie; klinische Diagnostik
			>200 ng/ml		38	1	132	169	22,4	99,4	
Nguyen, 2002 [M31]	USA	163 Patienten mit HCC+HCV, 149 Kontrollen mit HCV- Zirrhose	>20 ng/ml	312	103	30	60	119	63	79,9	Histologie; klinische Diagnostik
			>200 ng/ml		52	0	111	149	32,1	100	
Marrero, 2003 [M28]	USA	55 HCC-Patienten, 53 Kontrollen mit Zirrhose und Child- Pugh-Score <7	>20 ng/ml	108	37	10	18	43	68	82	Histologie
Lopez, 1996 [M25]	Malaysia	80 HCC-Patienten, 76 Kontrollen mit benigner Lebererkrankung	>200 ng/ml	156	47	2	33	74	58,8	97,4	Histologie
Okazaki, 1984 [M32]	Japan	Asymptomatische Patienten mit chronischer Lebererkrankung, AFP ↑	<20 ng/ml	245	11	44	3	187	78,6	80,9	Histologie; klinische Diagnostik
			<200 ng/ml		9	4	5	227	64,2	98,3	

Autor/Jahr	Land	Patientenkollektiv	AFP-Cut-off-Wert	N	RP	FP	FN	RN	SENS	SPEZ	Referenzstandard
Tong, 2001 [M43]	USA	Patienten mit Hepatitis B, C oder beide, an Klink überwiesen	>21 ng/ml	602	13	34	18	537	41	94	Histologie, Follow-up mind. 12 Monate
Chalasani, 1999 [M8]	USA	Patienten mit Zirrhose vor Transplantation	>20 ng/ml	285	17	34	10	224	63	87	Histologie; in einem Fall: typische Bildgebung
			>200 ng/ml		7	0	20	258	27	100	
Gambarin- Gelwan, 2000 [M14]	USA	Patienten mit Zirrhose vor Transplantation	>20 ng/ml	106	11	8	8	79	58	91	Leberexplantat

## **4.9 Ergebnisse der Meta-Analyse**

### **4.9.1 Ergebnisse für die Diagnostik mit US**

Die Ergebnisse der Analyse der 18 Studien mit Ultraschall sind in Tabelle 6 aufgeführt.

	<b>Sensitivität</b> <b>[95%-Konfidenzintervall]</b>	<b>Spezifität</b> <b>[95%-Konfidenzintervall]</b>
<b>Studien mit US</b>	66,2 [54,0;76,5]	95,7 [91,1;98,0]

**Tabelle 6:** Sensitivität und Spezifität für US

Aus der Studie *Tanaka et al.* ([M41]) gingen nur die Werte der Ultraschall-Analyse nach 6 Monaten in die Berechnung ein.

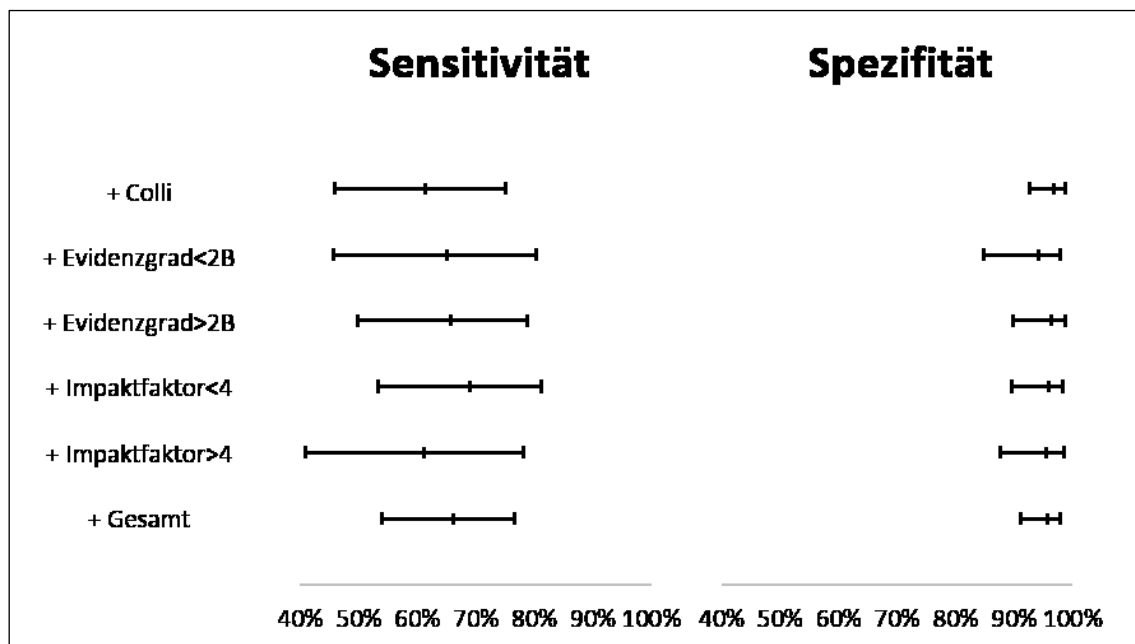
Die bivariate Analyse aller US-Studien ergab einen gepoolten Wert für die Sensitivität von 66,2% (KI 54,0% - 76,5%) und für die Spezifität von 95,7% (KI 91,1% - 98,0%). In der ersten Subgruppenanalyse, in die nur die Studien aus der Meta-Analyse von *Colli et al.* (2006) ausgewertet wurden, ergab sich eine im Vergleich zum Gesamtergebnis geringere Sensitivität sowie eine geringere Spezifität. Für die zweite und dritte Subgruppe, in die nur Studien mit einem Evidenzgrad  $\leq 2B$  bzw.  $> 2B$  aufgenommen wurden, wurden ebenfalls niedrigere Werte für die Sensitivität sowie für die Spezifität ermittelt. Der Vergleich der Subgruppen untereinander zeigte eine Überlegenheit der Studien mit Evidenzgrad  $> 2B$ , der in Bezug auf die Sensitivität nur minimal ist (65,8% vs. 65,0%), jedoch übertraf die Spezifität dieser Studien mit einem Wert von 96,4% deutlich diejenige der Studien mit höherem Evidenzgrad (Spezifität=94,2%). Der Sensitivitätswert der Studien aus der Meta-Analyse von *Colli et al.* lag mit 61,4% weit unter den Werten der beiden anderen Subgruppen, jedoch fand sich hier mit 96,9% die höchste Spezifität aller Untergruppen.

Die Ergebnisse der Untergruppen sind in Tabelle 7 zu sehen. Die Subgruppe *Lesion-by-lesion* beinhaltete bei den Ultraschallstudien nur eine einzelne Studie ([M43]), sodass hier keine eigene Analyse durchgeführt wurde.

	<b>Sensitivität</b> [95%-Konfidenzintervall]	<b>Spezifität</b> [95%-Konfidenzintervall]
<b>Studien aus der Meta-Analyse von Colli et al. (2006)</b>	61,4 [45,8;75,0]	96,9 [92,6;98,7]
<b>Studien mit einem Evidenzgrad Level 2B</b>	65,0 [45,8;80,4]	94,2 [84,7;97,9]
<b>Studien mit einem Evidenzgrad ab Level 3B</b>	65,8 [49,8;78,8]	96,4 [89,8;98,8]

**Tabelle 7:** Ergebnisse der Subgruppen-Analyse für US

Die Ergebnisse sind in Abbildung 8 graphisch als Forest plot dargestellt. Auf die Ergebnisse der beiden in dieser Grafik enthaltenen Subgruppen Impaktfaktor<4 und Impaktfaktor>4 wird im Abschnitt 4.10 eingegangen.



**Abb. 8:** Forest plots für Sensitivität und Spezifität der Ultraschall-Studien

#### 4.9.2 Ergebnisse für die Diagnostik mit CT

In Tabelle 8 sind die Ergebnisse aller diagnostischen Studien mit CT-Untersuchung aufgeführt. Die 13 in die bivariate Analyse aufgenommenen Studien ergaben einen Ergebniswert von 72,2% (KI 63,5% - 79,5%) für die gepoolte Sensitivität und einen Wert von 90,3% (KI 84,8% - 93,9%) für die gepoolte Spezifität.

	<b>Sensitivität</b> <b>[95%-Konfidenzintervall]</b>	<b>Spezifität</b> <b>[95%-Konfidenzintervall]</b>
<b>Studien mit CT</b>	72,2 [63,5;79,5]	90,3 [84,8;93,9]

**Tabelle 8:** Sensitivität und Spezifität für CT

Aus der Studie von *Talbot et al. (2010)* [M40] wurden die Werte für die Untersuchungen mit dem Kontrastmittel FDG ausgewählt. Ergebnisse von Untersuchungen mit anderen Kontrastmitteln aus dieser Studie wurden in die Gesamtanalyse nicht aufgenommen.

Die in den Untergruppen der CT-Studien errechneten Ergebnisse sind in Tabelle 9 dargestellt.

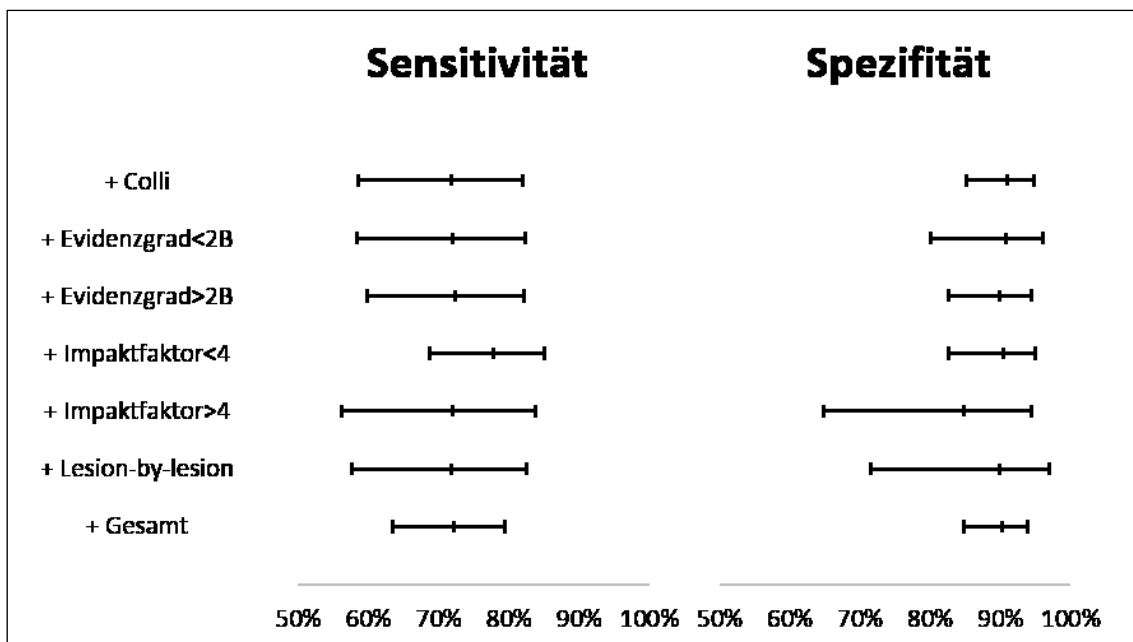
	<b>Sensitivität</b> <b>[95%-Konfidenzintervall]</b>	<b>Spezifität</b> <b>[95%-Konfidenzintervall]</b>
<b>Studien aus der Meta-Analyse von Colli et al. (2006)</b>	71,7 [58,6;82,0]	91,0 [85,2;94,7]
<b>Studien mit einem Evidenzgrad Level 2B</b>	71,9 [58,4;82,4]	90,8 [80,0;96,0]
<b>Studien mit einem Evidenzgrad ab Level 3B</b>	72,4 [59,8;82,1]	90,0 [82,6;94,4]
<b>Studien mit Lesion-by-lesion-Analyse</b>	71,7 [57,7;82,5]	90,0 [71,5;97,0]

**Tabelle 9:** Ergebnisse der Subgruppen-Analyse für CT



Anschließend wurden die Studien in 4 Subgruppen aufgeteilt und erneut analysiert. Die Ergebnisse für die Sensitivität unterschieden sich in den einzelnen Subgruppen nur gering. Am höchsten fiel die Sensitivität hier für die Studien mit einem Evidenzgrad schlechter als 2B aus (Sensitivität = 72,4%). Auch die errechneten Werte für die Spezifität in diesen Untergruppen variieren nicht stark; hier waren die Studien aus der Meta-Analyse von *Colli et al. (2006)* diejenigen mit der höchsten Spezifität (91,0%). Die niedrigsten Werte für die Spezifität ergaben sich für die Studien mit der Lesion-by-lesion-Analyse.

Einen graphischen Überblick über die CT-Subgruppen-Analyse liefert Abbildung 9. Auf die Ergebnisse der in dieser Grafik enthaltenen Subgruppen Impaktfaktor<4 und Impaktfaktor>4 geht der Abschnitt 4.10 ein.



**Abb. 9:** Forest plots für Sensitivität und Spezifität der CT-Studien

#### 4.9.3 Ergebnisse für die Diagnostik mit MRT

Tabelle 10 zeigt die Ergebnisse der Analyse aller eingeschlossenen MRT-Studien.

	<b>Sensitivität</b> <b>[95%-Konfidenzintervall]</b>	<b>Spezifität</b> <b>[95%-Konfidenzintervall]</b>
<b>Studien mit MRT</b>	82,0  [71,9;89,0]	90,3  [83,5;94,4]

**Tabelle 10:** Sensitivität und Spezifität für MRT

Die bivariate Analyse aus den Studien mit MRT als diagnostischem Test ergab für die zusammengefasste Sensitivität einen Wert von 82,0% (KI 71,9 – 89,0%) und für die zusammengefasste Spezifität einen Wert von 90,3% (KI 83,5% - 94,4%). Es folgte eine weitere Analyse in 4 Subgruppen.

Die Ergebnisse der Subgruppen-Analyse der MRT-Studien sind in Tabelle 11 aufgeführt.

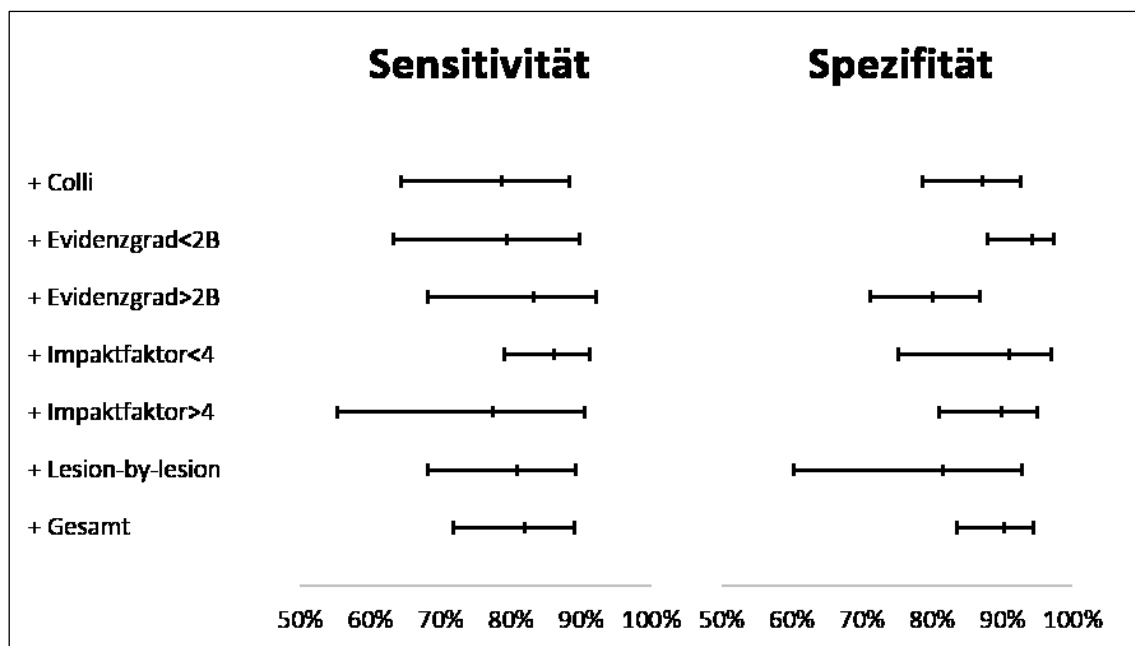
	<b>Sensitivität</b> <b>[95%-Konfidenzintervall]</b>	<b>Spezifität</b> <b>[95%-Konfidenzintervall]</b>
<b>Studien aus der Meta-Analyse von Colli et al. (2006)</b>	78,8 [64,4;88,4]	87,2 [78,7;92,6]
<b>Studien mit einem Evidenzgrad Level 2B</b>	79,5 [63,3;89,7]	94,2 [87,8;97,3]
<b>Studien mit einem Evidenzgrad ab Level 3B</b>	83,3 [68,2;92,1]	80,1 [71,2;86,7]
<b>Studien mit Lesion-by-lesion-Analyse</b>	80,9 [68,2;89,3]	81,6 [60,3;92,8]

**Tabelle 11:** Ergebnisse der Subgruppen-Analyse für MRT

Mit einem Wert von 83,3% für die Sensitivität war diese in der Subgruppe *Evidenzgrad schlechter 2B* am höchsten. Die niedrigste Sensitivität fand sich bei den Studien aus der Meta-Analyse von *Colli et al. (2006)*. Hier lag sie bei 78,8%.

Eine andere Reihenfolge ergab sich in der Spezifitäts-Analyse: Studien mit einem Evidenzgrad 2B hatten die höchste Spezifität (94,2%). Darauf folgten die Studien aus der Meta-Analyse von *Colli et al. (2006)* (Spezifität = 87,2%) und die Studien mit Lesion-by-lesion-Analyse (Spezifität = 81,6%), schließlich jene Studien mit einem Evidenzgrad >2B (Spezifität = 80,1%).

Abbildung 10 zeigt den Forest plot der MRT-Analysenergebnisse für Sensitivität und Spezifität. Im Abschnitt 4.10 werden die Ergebnisse der beiden in dieser Grafik enthaltenen Subgruppen Impaktfaktor<4 und Impaktfaktor>4 beschrieben.



**Abb. 10:** Forest plots für Sensitivität und Spezifität der MRT-Studien

#### 4.9.4 Ergebnisse für die Diagnostik mit AFP

Die Ergebnisse dieser Analyse sind in Tabelle 12 aufgeführt.

	<b>Sensitivität</b> <b>[95%-Konfidenzintervall]</b>	<b>Spezifität</b> <b>[95%-Konfidenzintervall]</b>
<b>AFP &gt; 20 ng/ml</b>	62,2 [53,5;70,2]	87,3 [82,2;91,1]
<b>AFP &gt; 200 ng/ml</b>	39,6 [23,9;57,8]	98,9 [97,9;99,4]

**Tabelle 12:** Sensitivität und Spezifität für AFP

Insgesamt gingen 8 Studien mit AFP als diagnostischem Test in die Analyse ein. Es wurden zwei Subgruppen analysiert, die aus 7 bzw. 5 Studien bestanden. Da die Studie von *Tong et al. (2001)* [M43] als einzige der AFP-Studien mit dem Evidenzgrad 2B versehen worden ist, wurden hierfür keine weiteren Subgruppen-Analysen durchgeführt. Das Ergebnis der bivariaten Analyse für die Subgruppe AFP > 20ng/ml ergab für die Sensitivität einen zusammengefassten Wert von 62,2% (KI 53,5% - 70,2%) und für die Spezifität einen zusammengefassten Wert von 87,3% (KI 82,2% - 91,1%).

Für Studien, die einen AFP-Wert von über 20 ng/ml als malignitätsverdächtig ansahen, fand sich eine höhere Sensitivität als für Studien, deren Cut-Off-Wert bei 200 ng/ml lag. Für die Spezifität dieser Studien kam genau das gegenteilige Verhältnis zustande: eine höhere Spezifität fand sich bei Studien mit einem AFP-Grenzwert von 200 ng/ml als bei Studien mit einem Wert von 20 ng/ml.

#### **4.10 Impaktfaktoren**

In zwei Fällen von älteren Studien, *Okazaki et al. (1984)* [M32] und *Lopez et al. (1996)* [M25] wurde der Impaktfaktor des Publikationsjahres nicht gefunden und stattdessen der 5-Jahres-Impaktfaktor benutzt. Alle anderen Studien wurden mit dem entsprechenden Impaktfaktor des Publikationsjahres versehen.

Tabelle 13 zeigt die Ergebnisse der US-Studien mit einem Impaktfaktor > 4, bzw. <4. Für die graphische Darstellung siehe Abbildung 8.

<b>US</b>	<b>Sensitivität [95%-Konfidenzintervall]</b>	<b>Spezifität [95%-Konfidenzintervall]</b>
<b>Impaktfaktor &lt; 4</b>	68,9 [53,2;81,2]	95,9 [89,6;98,4]
<b>Impaktfaktor &gt; 4</b>	61,1 [40,8;78,2]	95,6 [87,6;98,5]

**Tabelle 13:** Analyse-Ergebnisse für die US-Studien in Bezug auf den Impaktfaktor

Während die Spezifität in beiden Untergruppen kaum variiert (95,6% vs. 95,9%), lässt sich für die Sensitivität mit 68,9% ein höherer Wert für die Studien aus Zeitschriften mit niedrigem Impaktfaktor bestimmen. In die Untergruppe *Impaktfaktor > 4* wurden 6 Studien aufgenommen ([M13], [M19], [M23], [M37], [M42], [M48]). Einen kleineren Impaktfaktor hatten die übrigen 12 Studien.

Für die CT-Studien ergab sich ein ähnliches Bild. Auch hier war die Sensitivität in der Gruppe der Studien mit kleinem Impaktfaktor größer (77,9%), wohingegen die Spezifität in derselben Gruppe mit 90,5% deutlich höher lag als in der Gruppe mit großem Impaktfaktor (Spezifität = 84,8%). Eine Übersicht der Ergebnisse geben Tabelle 14 sowie Abbildung 9 wieder.

<b>CT</b>	<b>Sensitivität [95%-Konfidenzintervall]</b>	<b>Spezifität [95%-Konfidenzintervall]</b>
<b>Impaktfaktor &lt; 4</b>	77,9 [68,6;85,0]	90,5 [82,6;95,0]
<b>Impaktfaktor &gt; 4</b>	72,0 [56,2;83,8]	84,8 [64,9;94,4]

**Tabelle 14:** Analyse-Ergebnisse für die CT- Studien in Bezug auf den Impaktfaktor

Tabelle 15 zeigt die Ergebnisse der MRT-Studien. Für die graphische Darstellung siehe Abbildung 10. Hier fielen die Ergebnisse sowohl für die Sensitivität als auch für die Spezifität in der Gruppe der Studien mit niedrigem Impaktfaktor höher aus. Vor allem die Sensitivität unterschied sich mit 86,2% deutlich von der Sensitivität der Studien mit hohem Impaktfaktor, welche bei 77,4% lag.

<b>MRT</b>	<b>Sensitivität [95%-Konfidenzintervall]</b>	<b>Spezifität [95%-Konfidenzintervall]</b>
<b>Impaktfaktor &lt; 4</b>	86,2 [79,1;91,1]	90,9 [75,2;97,0]
<b>Impaktfaktor &gt; 4</b>	77,4 [55,3;90,5]	89,9 [81,0;94,9]

**Tabelle 15:** Analyse-Ergebnisse der MRT-Studien in Bezug auf den Impaktfaktor

Allgemein gaben die Ergebnisse der bivariaten Analyse für Studien mit einem Impaktfaktor <4 für die Subgruppe MRT die höchste Sensitivität mit einem Wert von 86,2% (KI 79,1% - 91,1%). Die niedrigste Sensitivität fand sich für Studien mit US als diagnostischem Test und Impaktfaktor > 4 (Sensitivität = 61,1%).

Die Spezifität hingegen war mit 95,9% (KI 89,6% - 98,4%) am höchsten in der Gruppe der Studien mit US als diagnostischem Test und niedrigem Impaktfaktor. Am niedrigsten war sie in der Gruppe der Studien mit CT als diagnostischem Test und hohem Impaktfaktor (Spezifität = 84,8%).

Der Zusammenhang zwischen Impaktfaktor und Sensitivität bzw. Spezifität wurde daraufhin noch mithilfe eines zweiseitigen t-Tests für unabhängige Stichproben untersucht. Hierbei zeigte sich für keine der bildgebenden Methoden ein signifikanter Zusammenhang zwischen der Höhe des Impaktfaktors der Fachzeitschrift und der jeweiligen Sensitivität bzw. Spezifität. Lediglich für die Sensitivität der CT-Untersuchung von Studien aus Zeitschriften mit einem Impaktfaktor > 4 wurde ein signifikantes Ergebnis nur knapp verpasst ( $p = 0,055$ ); die anderen Untersuchungen zeigten durchgehend nicht signifikante Ergebnisse ( $p$ -Werte zwischen 0,27 und 0,74).

## 5. DISKUSSION

### **5.1 Interpretation der Ergebnisse und Vergleich der Subgruppen-Analyse**

Die Ergebnisse der bivariaten Analyse bestätigen die diagnostischen Richtigkeiten der gängigen Bildgebung bei der Fragestellung „Verdacht auf HCC“ wie sie bisher angenommen wurden. Für die Sensitivität der Ultraschall-Untersuchungen ergab sich aus den Daten dieser Meta-Analyse ein Wert von 66,2% bei einer zugehörigen Spezifität von 95,7%. Damit ist die Ultraschall-Untersuchung in beiderlei Hinsicht der AFP-Messung überlegen, lediglich die Spezifität für einen Cut-Off-Wert von 200 ng/ml lag beim AFP mit 98,9% knapp über der des Ultraschalls. In der Subgruppen-Analyse waren die Studien mit einem Impaktfaktor kleiner 4 diejenigen mit der höchsten Sensitivität. Dass in Fachzeitschriften mit geringem Impaktfaktor nicht nur qualitativ schlechte sondern auch gut geplante und angemessen durchgeführte Studien veröffentlicht bzw. aufgenommen werden, zeigt sich durch die Betrachtung der Ergebnisse der Subgruppen *Evidenzgrad*. Hier fand sich ein nur minimaler Unterschied in den Ergebnissen für die Sensitivität zwischen Studien mit gutem Evidenzgrad  $\leq 2B$  bzw. mit schlechterem Evidenzgrad  $> 2B$  (65,0% vs. 65,8%). Die höchste Spezifität fand sich mit 96,4% in der Subgruppe der Studien mit schlechterem Evidenzgrad. Als Erklärung könnten Mängel in der Durchführung oder Fehler in der Auswahl des Patientenkollektivs dienen. Dies würde zu besseren Ergebnissen, aber einer schlechteren qualitativen Bewertung der Studien führen.

Die Ergebnisse für die CT-Untersuchungen zeigten mit einer Sensitivität von 72,2% (KI 63,5% - 79,5%) und einer Spezifität von 90,3% (KI 84,8% - 93,9%) eine gute diagnostische Richtigkeit. Diese konnte nur noch durch die MRT-Untersuchung übertroffen werden (Sensitivität = 82,0%, Spezifität = 90,3%). Genau wie bei den Subgruppen der US- und der MRT-Analyse fand sich in der CT-Subgruppe der Studien mit Impaktfaktor kleiner 4 die höchste Sensitivität mit einem Wert von 77,9%. Die Spezifität in dieser Untergruppe war mit 90,5% allerdings niedriger als diejenige der Subgruppe *Evidenzgrad*  $\leq 2B$  mit 90,8%.

Es stellt sich die Frage, ob in Fachzeitschriften mit hohem Impaktfaktor auch gleichzeitig nur qualitativ hochwertige Studien veröffentlicht werden. Der Impaktfaktor eines Journals ist immer auch davon abhängig, wie häufig dieses erscheint und wieviele Studien es beinhaltet. So wird ein Artikel aus einer Zeitschrift wie dem *New England Journal of Medicine*, welche wöchentlich und in englischer Sprache erscheint, mit einer höheren Wahrscheinlichkeit häufiger zitiert, als ein Artikel, der in einer Fachzeitschrift mit 12 Ausgaben im Jahr in nicht englischer Sprache veröffentlicht wurde. Diese Tatsache bleibt in der Angabe des Impaktfaktors unberücksichtigt. Eine Auswirkung auf die Höhe des Impaktfaktors haben außerdem die Anzahl der Forscher in einem bestimmten Fachgebiet. Ein Thema, an dem auf der ganzen Welt viel geforscht wird, wird mit großer Wahrscheinlichkeit öfter in Fachzeitschriften veröffentlicht, als seltene Themen, denen nur wenige Forscher weltweit ihre Aufmerksamkeit schenken. Auch dies führt dazu, dass solche Artikel häufiger zitiert werden und somit den Impaktfaktor einer Zeitschrift erhöhen können. In dieser Metaanalyse lieferten die Ergebnisse von Studien aus Fachzeitschriften mit hohem Impaktfaktor durchweg niedrigere Werte für Sensitivität und Spezifität als von Studien aus Journals mit niedrigerem Impaktfaktor. Dies kann so interpretiert werden, dass in Fachzeitschriften mit hohem Impaktfaktor, die folglich häufig zitiert werden, im Allgemeinen qualitativ hochwertigere Studien zu finden sind, welche eine niedrigere Sensitivität als Ergebnis haben und deren Ergebnisse somit näher an der Realität sind. Wohingegen Fachzeitschriften mit niedrigerem Impaktfaktor auch Studien von geringerer Qualität aufnehmen, was zu einer Überschätzung der Ergebnisse führen kann. Aus Mangel an Alternative [34] und weil der Impaktfaktor zwar nicht dafür geeignet ist, die Qualität von einzelnen Artikeln oder einzelnen Autoren zu bestimmen, wohl aber dafür, die Qualität einer Fachzeitschrift anzugeben [57], findet er weiterhin breite Anwendung. Da seine Aussagekraft über die Qualität der Inhalte der Fachzeitschrift jedoch eingeschränkt ist (siehe Kapitel 3.9 Vorbereitung für die Analyse – Impaktfaktor), sollte das Ergebnis in dieser Subgruppe kritisch betrachtet werden. Die hierfür zusätzlich durchgeführte Untersuchung mithilfe des t-Tests zeigte keine statistisch signifikanten Ergebnisse für den Zusammenhang zwischen Höhe des Impaktfaktors einer Zeitschrift und Höhe der Sensitivität bzw. Spezifität in der betreffenden Studie.



Die Subgruppe *Evidenzgrad*  $\leq 2B$  erzielte in der MRT-Analyse auch im Bezug auf die Spezifität das beste Ergebnis mit 94,2%, jedoch gingen nur 6 Studien in diese Subgruppe ein. Für die Subgruppe *Evidenzgrad*  $> 2B$  fand sich in der MRT-Analyse die niedrigste Spezifität mit 80,1%. Die Ergebnisse für die Sensitivität sind mit 83,3% relativ hoch einzuordnen.

Die niedrigste Sensitivität fand sich in der Subgruppe der Studien aus einer Fachzeitschrift mit Impaktfaktor  $> 4$ , was auf den ersten Blick eine höhere Qualität dieser Studien vermuten lässt, da gut geplante Studien oft ein realistischeres Bild zeichnen und somit zu niedrigeren Ergebnissen führen als Studien mit Fehlern in der Durchführung. Tatsächlich findet sich der zweitniedrigste Wert für die Sensitivität in der Untergruppe der Studien mit hoher Evidenz, sodass ein solcher Schluss nahe liegt.

Die geringste Sensitivität ergab sich für die Messung des AFP-Werts, wobei diese noch erheblichen Schwankungen unterlag, je nach dem wo der Cut-Off-Wert gesetzt wurde. So schwankte dieser Wert in dieser Meta-Analyse zwischen 39,6% für einen Cut-off-Wert von 200 ng/ml und 62,2% bei einem Grenzwert von 20 ng/ml. Der Zugewinn an Sensitivität für diesen Wert geht allerdings mit einem Verlust an Spezifität einher (98,9% bzw. 87,3%). Dies liegt daran, dass der AFP-Wert ebenfalls als Tumormarker für Tumoren embryonalen Ursprungs gilt [5]. Dies bedeutet, dass auch Keimzelltumoren zu einem Anstieg des AFP-Wertes führen; allerdings sind besonders hohe AFP-Werte typisch für HCC, sodass bei höheren Grenzwerten die Spezifität sehr hoch liegt, die Sensitivität jedoch eher niedrig ist [9]. Die Ergebnisse dieser Studie stützen diese Aussage.

Allgemein lässt sich sagen, dass die hier zugrunde liegenden Studien z.T. lediglich einen geringen Evidenzgrad haben und somit qualitative Mängel aufweisen. Das Fehlen von qualitativ höher wertigen Studien hat somit auch eine Auswirkung auf das Ergebnis der Meta-Analyse, welche allerdings durch die Summe der einzelnen Studien trotzdem einen guten Überblick verschaffen kann.

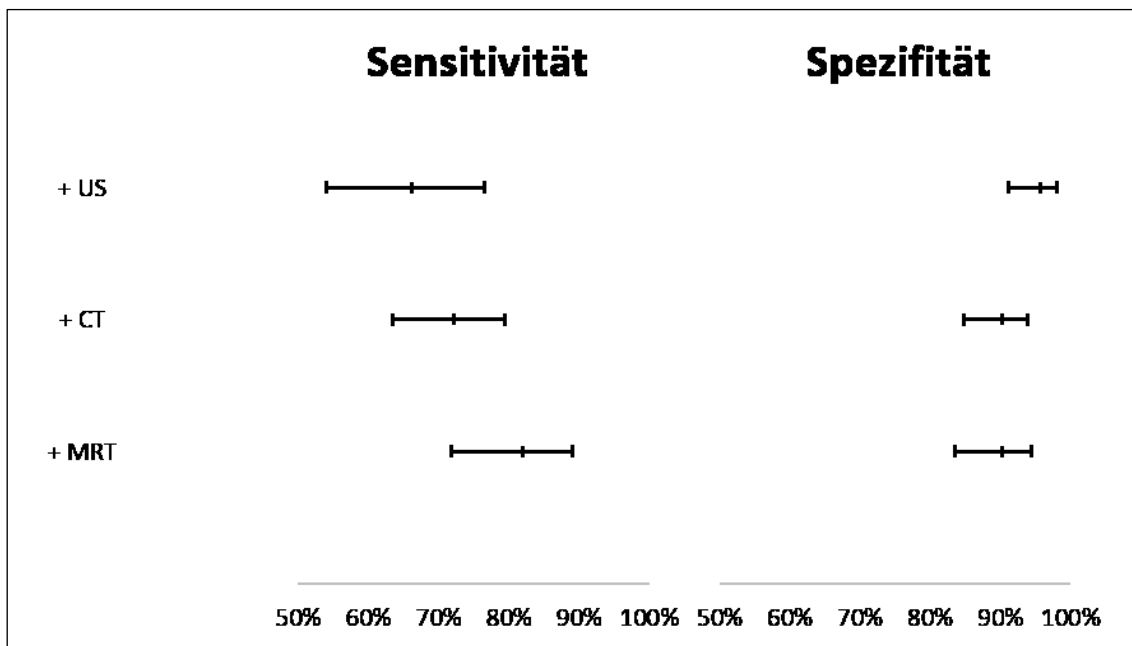
Ein allgemeines Problem in der Diagnostik maligner Erkrankungen ist die Frage nach dem Zeitpunkt der Diagnosestellung. Durch immer neue Untersuchungs- und Bildgebungsmethoden und Fortschritte in der Ausführung von Screeningprogrammen ist es heutzutage möglich, bestimmte Erkrankungen immer früher zu entdecken.

Allerdings verfälscht sich dadurch mitunter die Wahrnehmung in Bezug auf die Lebensverlängerung der Patienten. Ein fiktives Beispiel: Noch vor 10 Jahren wurden Karzinome durch eine bestimmte Bildgebung erst im Endstadium diagnostiziert und die Patienten lebten im Durchschnitt nach der Diagnosestellung noch ein Jahr. Das bewusste Leben mit der Krankheit betrug somit etwa ein Jahr. Heute jedoch sei es möglich, diese Karzinome bereits 5 Jahre vorher zu entdecken. Die Patienten leben also insgesamt 6 Jahre mit der Krankheit, sterben aber zur gleichen Zeit. So kann der Eindruck entstehen, dass die Patienten durch die verbesserte Diagnostik fünf Jahr länger gelebt hätten als zuvor, wobei sie jedoch nur länger in dem Bewusstsein der Erkrankung gelebt haben. Dieses Phänomen, bei dem nur eine scheinbare Lebenszeitverlängerung des Patienten aufgrund einer vorverlegten Diagnosestellung erreicht wird, wird als *lead-time bias* bezeichnet [21].

## **5.2 Fazit**

Insgesamt bestätigt diese Studie zwar die Überlegenheit der diagnostischen Richtigkeit von CT und MRT gegenüber US und AFP, für den klinischen Alltag ergeben sich daraus allerdings keine Konsequenzen. Das HCC-Screening Patienten einer Risikogruppe muss flächendeckend verfügbar, kostengünstig, wenig invasiv und schnell durchführbar sein um eine große Anzahl von Patienten in regelmäßigen Abständen untersuchen zu können. Die CT-Untersuchung als solche birgt für den Patienten eine nicht unerhebliche Strahlenbelastung, sodass sie v.a. zur Charakterisierung und OP-Planung eingesetzt werden sollte. Die MRT-Untersuchung hat zusammen mit der CT-Untersuchung das Problem, dass sie zeitaufwendig und teuer ist, außerdem sind diese Geräte nicht an jedem Krankenhaus zu finden. Dies führt dazu, dass der Einsatz von CT und MRT in der Diagnostik des hepatozellulären Karzinoms gut durchdacht und zielgerichtet sein sollte.

Abbildung 11 zeigt die Sensitivitäten und die Spezifitäten der wichtigsten diagnostischen Tests in der Diagnostik des hepatozellulären Karzinoms mit ihren zugehörigen Konfidenzintervallen.



**Abb. 11:** Sensitivität und Spezifität von US, CT und MRT

### **5.3 Vergleich der Ergebnisse mit der Analyse von Colli et al. (2006)**

Die bivariate Analyse aus dieser Studie ergab für die US-Studien eine gepoolte Sensitivität von 66,2% (KI 54,0% - 76,5%) und eine gepoolte Spezifität von 95,7% (KI 91,1% - 98,4%). Im Vergleich hierzu lagen die Ergebnisse der Studie von *Colli et al.* bei 60% (KI 44% - 76%) für die Sensitivität, bzw. 97% (KI 95% - 98%) für die Spezifität.

Für die CT-Studien fand sich hier folgendes Ergebnis: Sensitivität 72,2% (KI 63,5% - 79,5%), Spezifität 90,3% (KI 84,8% - 93,9%). Die Autoren der Studie selber errechneten für die Sensitivität einen Wert von 68% (KI 55% - 80%) und für die Spezifität von 93% (KI 89% - 96%). Diese Unterschiede können dadurch zustande kommen, dass die Autoren *Colli et al.* eine Studie ([M6]) ohne Angabe von Gründen nicht in die CT-Analyse mit aufgenommen haben. Außerdem unterschieden sich die errechneten Werte der Sensitivität in zwei Studien ([M8], [M30]); da *Colli et al.* nur das Ergebnis für die Sensitivität wiedergeben und nicht die genauen Werte für RP, FP, RN und FN, kann keine genaue Aussage gemacht werden, worin sich die in dieser Meta-Analyse gefundenen Werte von denen von *Colli et al.* unterscheiden.

Die Daten der MRT-Studien ergaben in dieser Meta-Analyse einen Sensitivitäts-Wert von 82,0% (KI 71,9% - 89,0%) und einen Spezifitäts-Wert von 90,3% (KI 83,5% - 94,4%).

*Colli et al. (2006)* berechneten hier für die Sensitivität 81% (KI 70% - 91%) und für die Spezifität 85% (KI 77% - 93%).

Es zeigt sich also, dass in dem speziellen Fall der in der Meta-Analyse von *Colli et al.* verwendeten Studien nur geringe Unterschiede zwischen den Ergebnissen der Originalarbeit und dieser Dissertation bestehen. Die Werte unterscheiden sich ca. zwischen 1 und 6 %. Die eigene Analyse fand im Allgemeinen etwas bessere Werte für die Sensitivität und etwas schlechtere Werte für die Spezifität als die Autoren der Meta-Analyse *Colli et al.* Nur im Falle der MRT-Studien fanden sich sowohl für die Sensitivität als auch für die Spezifität bessere Werte in der eigenen Analyse. Allerdings ist das bivariate Modell hier die angemessenere Analysemethode, weshalb es in der eigenen Meta-Analyse Verwendung fand.

Die bestehenden Unterschiede können zum einen zustande kommen durch die Verwendung des bivariaten Modells in der eigenen Analyse, wohingegen *Colli et al.* eine andere Art der Auswertung verwendeten (summary receiver operating characteristic, SROC). Andererseits wurden bei Diskrepanzen zwischen Daten, die in der Meta-Analyse von *Colli et al.* angegeben wurden, und den Daten aus der Originalstudie letztere in die eigene Analyse aufgenommen. Außerdem wurden in der eigenen Analyse Studien mit Patient-by-patient- und Lesion-by-lesion-Angabe getrennt bewertet, wohingegen bei *Colli et al.* keine Unterscheidung stattfand. Somit fielen in der eigenen Analyse M36 und M49 aus der Gesamtbewertung heraus. Trotzdem haben diese Abweichungen zu keinem auffallenden Ergebnisunterschied geführt.

#### **5.4 Vergleich der diagnostischen Richtigkeit mit Ergebnissen anderer Autoren**

Unterschiede in der Auswahl der gefundenen Studien lassen sich für die Studie von *Guang et al. (2011)* [36] v.a. durch die im Suchvorgang von dieser Dissertation abweichenden Suchbegriffe wie etwa „SonoVue®“, „hepatic adenoma“ oder „intrahepatic cholangiocarcinoma“ erklären. Der Vergleich mit den Ergebnissen von *Guang et al. (2011)* muss immer mit dem Bewusstsein geschehen, dass hier auch andere Leberläsionen als HCC als Zielkrankheit definiert waren.

*Guang et al. (2011)* errechneten hier für die Sensitivität des US einen Wert von 88% (KI 87% – 90%) und für die Spezifität von 81% (KI 79% – 84%). Diese Ergebnisse unterscheiden sich in beiden Werten relativ deutlich von den eigenen (Sensitivität = 66,2%, Spezifität = 95,7%).

*Guang et al. (2011)* kamen hier zu einem Ergebnis von 90% (KI 8% - 92%) für die CT-Sensitivität und 77% (KI 71% – 82%) für die CT-Spezifität. In der eigenen Dissertation fand sich für die CT-Untersuchung ein Sensitivitätswert von 72,2% (KI 63,5% - 79,5%) bei einer Spezifität von 90,3% (KI 84,8% - 93,9%).

*Guang et al. (2011)* fanden für das MRT einen gepoolten Sensitivitätswert von 86% (KI 83% - 88%) bei einer gepoolten Spezifität von 81% (KI 76% - 85%). Mit einer Sensitivität von 82,0% (KI 71,9% - 89,0%) ist das Ergebnis dieser Dissertation sehr nahe bei denen der anderen Autoren, wohingegen die Spezifität mit 90,3% (KI 83,5% - 94,4%) etwas höher liegt.

Ein Grund für die v.a. bei den US-Analysen gefundenen Unterschiede könnte die unterschiedliche Studiena Auswahl sein. Keine der eigenen Studien fand sich in der Analyse von *Guang et al.* und ebenso fand sich auch keine von deren Studien in der eigenen Analyse. Ein weiterer Unterschied zwischen der eigenen und der Meta-Analyse von *Guang et al.* liegt in der unterschiedlichen Methodik. Hier ist das bivariate Modell in der eigenen Analyse als besser geeignet anzusehen, um die Beziehung zwischen Sensitivität und Spezifität realistischer wiederzugeben. Außerdem waren in die Meta-Analyse der anderen Autoren weniger Studien eingeschlossen, nämlich 21. Für die einzelnen diagnostischen Tests lagen für die CEUS-Untersuchung 10 Studien vor (18 in der eigenen), für die CT-Untersuchung 9 (14 in der eigenen) und für die MRT-Untersuchung wieder 10 (13 in der eigenen Studie). In der eigenen Analyse wurden mit 48 Artikeln mehr als doppelt so viele Studien eingeschlossen, was ebenfalls eine Auswirkung auf die Ergebnisse haben kann. Ebenso fand sich bei *Guang et al.* das Fehlen eines Kontrastmittels in der CT oder MRT-Untersuchung als expliziter Ausschlussgrund, was in der eigenen Analyse nicht der Fall war. Ein weiterer Unterschied liegt in der unterschiedlichen Verwendung von Kontrastmitteln in den einzelnen Studien. In der Analyse von *Guang et al.* wurden für jeden diagnostischen Test sehr einheitliche Kontrastmittel gewählt, unverstärkte diagnostische Tests galten

als Ausschlusskriterium. In der eigenen Meta-Analyse wurden alle US-, CT- oder MRT-Techniken eingeschlossen, sodass hier eine heterogene Gruppe von Untersuchungsmethoden zusammengefasst wurde. Auf die gepoolten Werte des diagnostischen Tests kann das einen eher negativen Effekt gehabt haben, sodass tendenziell schlechtere Werte errechnet wurden.

Im Oktober 2013, während der letzten Überarbeitungen an dieser Dissertation, erschien in der Zeitschrift „Ultraschall in der Medizin“ ein Artikel, der inhaltlich sehr ähnlich ist. Die Autoren *Floriani et al.* [29] untersuchten ebenfalls mit Hilfe der bivariaten logistischen Regression die Daten verschiedener Studien zum Thema Diagnostik des hepatozellulären Karzinoms. Auf Grund der Aktualität dieses Artikels und der Nähe zum eigenen Thema, soll auf diese Meta-Analyse ein besonderes Augenmerk gerichtet werden.

*Floriani et al.* (2013) [29] suchten in den Datenbanken PubMed und EMBASE mit folgenden MeSH- und Volltext-Begriffen: **(hepatocellular carcinoma) AND (MRI AND MR OR CT OR SCT OR MSCT OR MDCT OR diagnostic tests OR US OR IOUS) AND (humans)**. Die Suche unterlag keinen Einschränkungen in Bezug auf die Sprache. Der Zeitraum der Veröffentlichung der Artikel reichte vom 01.01.1996 bis zum 30.06.2011. Die Ein- und Ausschlusskriterien für die Studien waren wie folgt: Angaben über richtig positive, richtig negative, falsch positive und falsch negative Werte oder die Möglichkeit, diese aus der Originalpublikation zu entnehmen, um Sensitivität und Spezifität zu errechnen; histopathologische Untersuchung als Referenzstandard; keine sequentiellen Tests; US, CT und/oder MRT als diagnostischer Test. Somit schlossen sie in ihre Analyse insgesamt 81 Studien ein. Hierbei unterschieden *Floriani et al.* zwischen Studien, die nur eine der interessierenden diagnostischen Tests verwendeten (n=16) und solchen, die mehr als einen dieser Tests benutzten (n=65). Die Ergebnisse wurden ebenfalls in Form einer gepoolten Sensitivität und gepoolten Spezifität mit jeweiliger Angabe der 95%-Konfidenzintervalle dargestellt.

*Floriani et al.* (2013) fanden für den Ultraschall einen gepoolten Wert für die Sensitivität von 86% (KI 82% - 90%) und für die Spezifität von 77% (KI 65% - 89%). Dieser unterscheidet sich recht deutlich von den in der eigenen Analyse errechneten Werten (Sensitivität = 66,2%, Spezifität = 95,7%). Als Gründe hierfür könnte die in den

Studien von *Floriani et al.* stärkere Berücksichtigung von Studien mit Lesion-by-lesion-Analyse sein. 62 der 81 Studien gaben ihre Ergebnisse nur in dieser oder als Segment-Analyse wieder. In der eigenen Dissertation hingegen wurden die Lesion-by-lesion-Daten nicht in die Hauptanalyse mit einbezogen. Die dadurch niedrigere Sensitivität könnte daher rühren, dass bei Patienten mit mehreren HCC-Läsionen die Entdeckung einer dieser Knoten wahrscheinlicher ist als bei einem Patienten mit nur einem Knoten, sodass die Patient-by-patient-Angabe hier schlechter ausfällt.

Für die Computertomographie berechneten *Floriani et al. (2013)* einen Sensitivitätswert von 78% (KI 74% - 81%) und einen Spezifitätswert von 77% (KI 54% - 99%). Die eigenen Werte für diese Bildgebung waren eine Sensitivität von 72,2% (KI 63,5% - 79,5%) und eine Spezifität von 90,3% (KI 84,8% - 93,9%). Hier liegen die Werte beider Meta-Analysen zumindest für die Sensitivität näher bei einander. Auch die MRT-Analyse von *Floriani et al. (2013)* ergab mit einem Wert für die zusammengefasste Sensitivität von 84% (KI 81% - 87%) und für die zusammengefasste Spezifität von ebenfalls 84% (KI 80% - 88%) der eigenen Analyse noch ähnlichere Werte (MRT-Sensitivität = 82,0%, Spezifität = 90,3%).

Weitere Gründe für Unterschiede in den Ergebnissen können die abweichenden beinhalteten Studien sein. Ursächlich hierfür war einerseits der Suchlogarithmus, der sich in beiden Fällen unterschiedlich darstellt. Die Literatursuche von *Floriani et al.* wurde mit einem relativ kurzen Suchbegriff und allgemein gehaltenen Suchbegriff mit weniger Kombinationsmöglichkeiten durchgeführt, sodass hier eine deutlich höhere Anzahl an Treffern zu verzeichnen war (> 5000 Treffer; im Vergleich: ca. 1000 Treffer in der eigenen Dissertation). Jedoch führte die offenere Suchanfrage wohl auch zum vermehrten Ausschluss von Studien, die in der Fragestellung nicht zum Thema passten. In die Volltextanalyse wurden 190 Studien aufgenommen, in der eigenen Dissertation waren es 165. Die ursprüngliche Masse an Treffern führte demnach nicht zu einer wesentlich größeren Anzahl an passenden Studien; sie könnte jedoch durchaus zu einer unterschiedlichen Auswahl an Studien beigetragen haben. Zudem konnten *Floriani et al.* auch italienische und (durch die Mitarbeit eines asiatischen Kollegen) chinesische Studien mit aufnehmen. Einige der hier aufgenommenen Studien wären durch fehlende Angabe von Spezifität von der eigenen Analyse ausgeschlossen worden. Zudem wählten *Floriani et al.* einen etwas anderen Veröffentlichungs-

zeitraum, nämlich vom Januar 1996 bis Juni 2011. In die eigene Analyse gingen Studien von Januar 2000 bis März 2011 (Erstsuchvorgang) bzw. Mai 2013 (Zweitdurchgang mit anderen Suchbegriffen) ein.

Als mögliche Ursache für die Unterschiede der enthaltenen Studien kommen außerdem die Ein- und Ausschlusskriterien in Frage. So konnten in die eigene Analyse auch Studien aufgenommen werden, die anstelle eines histopathologischen Referenzstandards auf Grundlage der EASL-Kriterien eine HCC-Diagnose stellen. Dies ist im klinischen Alltag häufig der Fall und wurde deshalb als Referenzstandard akzeptiert. Zwar wäre die alleinige histopathologische Bestätigung der Diagnose ideal, jedoch lässt sie sich aus verschiedenen Gründen nicht immer durchführen. So kann eine erhöhte Blutungsneigung oder eine lebensbedrohliche Komorbidität die Leberpunktion bzw. Lebertransplantation unmöglich machen. Für die adäquate Therapie ist eine Diagnosestellung aber unumgänglich, sodass hierfür die EASL-Kriterien eine Alternative zur Histopathologie bieten.

18 Studien sind sowohl in die Meta-Analyse von *Floriani et al. (2013)* als auch in die eigene Meta-Analyse eingegangen. 4 der eigenen Studien wurden von deren Analyse ausgeschlossen mit folgender Begründung: unsuitable reference standard (n=2, [M8], [M48]), unsuitable imaging test (n=1, [M37]), unsuitable experimental design (n=1, [M23]). Im Falle von *Chalasani et al. [M8]* handelt es sich um die Histologie von Biopsaten in 26 Fällen und der typischen CT- und Angiographie-Untersuchung in einem Fall. Bei *Yu et al. [M48]* ist der Referenzstandard die Korrelation im Leberexplantat. *Saada et al. [M37]* wurde von *Floriani et al.* ausgeschlossen, weil hier ein intraarterielles Kontrastmittel gegeben wurde, abweichend von den intravenösen Kontrastmitteln der anderen Studien. Die Studie von *Libbrecht et al. [M23]* schließlich untersuchte die explantierten Lebern von Zirrhotikern mit Vergleich der präoperativ erhaltenen Bilder. Hier wurden sowohl HCC als auch andere Leberläsionen beachtet, die Bewertung wurde aber getrennt angegeben.

In die eigene Analyse wurden schon nach Lesen der Abstracts einige der Studien aus *Floriani et al.* nicht übernommen, weil sie keine Angaben zur Spezifität machten (n=7), einen sequentiellen Test enthielten (n=1) oder die HCC-Charakterisierung zum Ziel hatten und nicht die Diagnostik (n=1).



Ein weiterer Unterschied im Vorgehen war die Bewertung der aufgenommen Studien. Die Autoren der italienischen Publikation verwendeten hierfür eine modifizierte Version des sogenannten *Quality Assessment of Diagnostic Accuracy Studies (QUADAS)* Tools. In der eigenen Meta-Analyse kam der Evaluationsbogen des *Oxford Centre for Evidence-based Medicine* (siehe Anhang 8.4 Anlage 4) zur Anwendung.

Auch in der Auswertung der sich in beiden Meta-Analysen überschneidenden Studien fanden sich Unterschiede. Für die CT-Ergebnisse von *Teefey et al.* [M42] wurde in der vorliegenden Meta-Analyse ein falsch-negativer Wert von 3 angegeben, die Autoren der Analyse aus dem Jahr 2013 fanden hier einen Wert von 5. Auch die Werte für die US- und die MRT-Untersuchungen weichen jeweils um 1 voneinander ab, da in dieser Studie 2 Untersucher für die Auswertung des diagnostischen Tests verantwortlich waren. In der eigenen Analyse wurde hierfür immer der Mittelwert beider Untersucher genommen und entsprechend auf- bzw. abgerundet. Die Gesamtpatientenzahl wurde durch die entsprechende Abänderung berücksichtigt, sodass immer eine Gesamtpatientenzahl von 25 beibehalten wurde. In der Meta-Analyse von *Floriani et al.* wurden auch entsprechende Auf- bzw. Abrundungen durchgeführt, jedoch schwankt hier die Gesamtpatientenzahl zwischen 24 und 26. Im Vergleich von US und CT sind für die Studie von *Teefey et al.* [M42] fälschlicherweise die MRT-Werte anstelle der US-Werte angegeben. Ob dies in der Berechnung einen Fehler darstellt oder nur ein Druck- oder Kopierfehler ist, kann nicht nachvollzogen werden.

Für die Studie von *De Lédinghen et al.* [M11] wurden in der eigenen Dissertation die Patient-by-Patient-Werte errechnet und ausgewertet. *Floriani et al.* wählten hierfür die Lesion-by-lesion-Angaben. Gleiches gilt für die Werte der Studie von *Burrel et al.* [M7].

Dieselben Werte wurden in beiden Meta-Analysen für die Studie *Rode et al.* [M36] angegeben.

Die Ähnlichkeit der beiden Studien zeigt, dass die Diagnostik des hepatozellulären Karzinoms ein sehr aktuelles Thema ist und weiterhin eine Herausforderung im klinischen Alltag darstellt, für die es gilt, gute Hilfestellungen zu finden. Es zeigt sich auch, dass das statistische Modell der bivariaten logistischen Regression gegenüber der ROC-Methode einen Fortschritt bringt und der somit stets berücksichtigte

zweidimensionale Aspekt der Sensitivität und Spezifität von diagnostischen Tests eine große Rolle spielt.

## **5.5 Potentielle Limitationen dieser Meta-Analyse**

Wie bei jeder wissenschaftlichen Arbeit können auch bei sorgfältigstem Vorgehen unterschiedliche Arten von Verzerrungen nicht mit absoluter Sicherheit ausgeschlossen werden.

Schon durch den Bewertungsvorgang geht eine gewisse Variation in die Ergebnisse ein. Zwar soll die Verwendung von standardisierten Fragebögen, wie z.B. dem hier verwendeten Fragebogen des Oxford Centre of Medicine, die subjektive Einflussnahme der Untersucher möglichst gering halten; jedoch kann eine völlig objektive Bewertung nie ganz erreicht werden. Einer potentiellen Verzerrung der Ergebnisse, welche Studien als relevant angesehen wurden bzw. mit welchem Evidenzgrad sie versehen wurden, wurde durch die Zusammenarbeit zweier Untersucher versucht entgegen zu wirken. So wurde bei Meinungsverschiedenheiten durch Diskussion ein Konsens getroffen. Jedoch besteht die Möglichkeit, dass bei anderen Untersuchern ein anderer Score bzw. eine andere Auswahl getroffen worden wäre.

Im Falle einer Meta-Analyse birgt schon der Suchvorgang bei der Literaturrecherche solche Verzerrungsquellen, die es zu berücksichtigen gilt. In dem konkreten Fall dieser Dissertation wurden im Suchvorgang nur Studien in deutscher oder englischer Sprache berücksichtigt. Viele Studien, die in anderssprachigen Ländern publiziert wurden, konnten somit keine Beachtung finden; dies gilt im besonderen Maße für den asiatischen Raum, der durch die erhöhte Inzidenz des HCC prädestiniert ist für einen größeren Umfang an Studien zu diesem Thema. Da allerdings häufig nur der Abstract einer Studie auf Englisch übersetzt wird, konnten alle Studien, die im Volltext eine andere Sprache als Englisch oder Deutsch aufwiesen, nicht zu einer Aufnahme in dieser Meta-Analyse führen.

Auch die Erstausswahl nach den Inhalten der Abstracts der Studien kann problematisch sein. Wurden wichtige Charakteristika der Studien nicht im Abstract erwähnt und waren diese Bestandteil der Ein- und Ausschlusskriterien dieser Dissertation, so konnte dies zum Ausschluss der Studien führen. Im Zweifelsfall wurde zwar der Volltext der

Studie besorgt und die Studie dann nach dessen Durchsicht bewertet, jedoch konnte eine geringe Qualität des Abstracts, der z.B. nur einen allgemeinen Überblick zum Thema gab, dazu führen, dass auch der Volltext der Studie nicht beachtet wurde. So könnte eine gute Studie bei lückenhafter Publikation als nicht geeignet angesehen worden sein, obwohl sie alle Kriterien zum Einschluss in die Meta-Analyse erfüllte.

Für die Literatursuche wurde im Verlauf dieser Dissertation die Datenbank PubMed benutzt. Eine zusätzliche Suche in einer weiteren Datenbank hätte eventuell weitere interessierende Publikationen für diese Meta-Analyse finden können. Da der Umfang von PubMed über 20 Millionen Zitate beinhaltet, die z.T. auch aus anderen Datenbanken wie etwa MEDLINE zustande kommen, ist jedoch nicht von einer allzu großen Masse nicht gefundener Studien auszugehen.

Zu einem sogenannten Publikationsbias kann es dennoch kommen, da in den genannten Datenbanken bestimmte Publikationen keinen Eingang finden. So können Studienergebnisse, die auf Fachkongressen vorgetragen und nicht publiziert werden oder in einer Klinik intern veröffentlicht werden, nicht in solch einer Datenbank gefunden werden. Auch ist die Wahrscheinlichkeit veröffentlicht zu werden für eine Studie, die keine signifikanten Ergebnisse liefert, deutlich geringer als für eine Studie, die ein klares Ergebnis zeigt [49].

Wurden in einer Originalpublikation zu bestimmten Fragestellungen, die für die Erhebung des Evidenzgrades von Bedeutung waren, keine genauen Angaben gemacht, so wurde in der Bewertung davon ausgegangen, dass diese Fragestellung nicht erfüllt wurde. Z.B. wurde davon ausgegangen, dass der Referenzstandard nicht unabhängig war, wenn nicht ausdrücklich im Text beschrieben wurde, dass der Referenzstandard ohne das Wissen um das Ergebnis im diagnostischen Test durchgeführt wurde. Das bedeutet, dass eine Studie aufgrund einer ungenauen Beschreibung der Abläufe einen schlechteren Evidenzlevel zugeteilt bekommen konnte, obwohl sie in Wahrheit alle Kriterien für eine bessere Bewertung erfüllte. Diese Annahme des „worst case scenario“ soll eine Überbewertung der Studienqualität verhindern, kann aber unter Umständen dazu führen, dass die Qualität der Studie als zu schlecht eingestuft wird. Gleiches gilt auch für die Auswahl der Ein- und Ausschlusskriterien. Um hier möglichst gut zur Fragestellung passende und aussagekräftige Studien zu finden, wurden

bestimmte Kriterien festgelegt, die die Studien erfüllen mussten. Dadurch wurden z.B. Studien, die vor dem Jahr 2000 veröffentlicht wurden, nicht in die Meta-Analyse mit aufgenommen. Außerdem wurden die Ein- oder Ausschlusskriterien bei Studien, die diese nicht explizit erwähnten, als nicht erfüllt angesehen. Eine ungenaue Angabe des Referenzstandards konnte somit zum Ausschluss der Studie führen.

Für die Vergleichbarkeit der Studienpopulationen ist zu sagen, dass zwar durch die Fokussierung auf die Risikopopulation in den einzelnen Studien jeweils Patienten mit ähnlicher Grunderkrankung aufgenommen wurden (v.a. Leberzirrhose). Da die Ätiologie für diese Grunderkrankung aber sehr unterschiedlich sein kann und epidemiologische und genetische Faktoren hier mit einfließen, ist die Vergleichbarkeit der Studienpopulationen nur bedingt gegeben. Gerade die Unterschiede zwischen dem asiatischen und dem westlichen Raum sind hierbei zu beachten.

Schließlich kann über die Methodik dieser Meta-Analyse gesagt werden, dass hier eine bivariate logistische Regression durchgeführt wurde, obwohl die häufigste Analysemethode diagnostischer Studien mit der paarweisen Angabe von Sensitivität und Spezifität das sogenannte „summary Receiver Operating Characteristic (sROC)“ ist [62]. Die Gründe hierfür sind im Kapitel *3.10 Bivariate logistische Regression* aufgeführt. Ein Vergleich der Ergebnisse dieser Dissertation mit denen anderer Autoren, die die weit verbreitete sROC-Methode anwenden, wird dadurch erschwert.

Für die Ergebnisse der verschiedenen Subgruppen-Analyse ist zu beachten, dass die einzelnen Untergruppen teilweise nur wenige Studien beinhalteten. So fanden sich in einigen Subgruppen nur 5 oder 6 Studien, was die Aussagekraft des Ergebnisses dieser Subgruppen-Analyse begrenzt. Um hier nicht zielführende Berechnungen zu vermeiden, wurde für eine Subgruppen-Analyse eine Mindestanzahl von 3 Studien vorausgesetzt.

## 6. ZUSAMMENFASSUNG

Mit über 600 000 Todesfällen pro Jahr ist der Leberkrebs weltweit unter den 10 häufigsten Krebstodesursachen zu finden [59]. Der mit bis zu 85% häufigste histologische Subtyp ist das hepatozelluläre Karzinom, das somit bereits im Jahr 1990 für über 5% aller Krebsfälle weltweit verantwortlich war [12]. Das HCC ist mit einer 5-Jahres-Überlebensrate von 6,5% mit einer sehr schlechten Prognose assoziiert [25]. In den letzten Jahren fand sich eine stetige Zunahme der HCC-Inzidenz v.a. in Ländern mit geringer HCC-Prävalenz wie den USA und Westeuropa. Hingegen wurde in Ländern mit hoher HCC-Prävalenz, wie in Teilen Asiens und Afrikas, ein Rückgang der Neuerkrankungen verzeichnet [27]. Weltweit sind Männer etwa doppelt so häufig betroffen wie Frauen [40]. Als Hauptrisikofaktor für die Entstehung eines HCC gilt die Leberzirrhose, wobei weltweit v.a. die Infektion mit dem Hepatitis-C-Virus eine hohe Inzidenz an HCC-Erkrankungen bei Patienten mit Zirrhose aufweist [28]. Eine frühe Diagnosestellung ist bei den geringen Therapieoptionen für die Patienten prognosebestimmend.

Ziel dieser Dissertation war es, die aktuelle Studienlage zur Diagnostik des HCC darzustellen, vorhandene Studien in ihrer Durchführung und Qualität zu bewerten und ihren Evidenzgrad zu ermitteln, um dann die Ergebnisse in Form einer Metaanalyse quantitativ zusammenzufassen. Hierfür wurde auf PubMed eine umfassende und systematische Literaturrecherche durchgeführt, die so gefundenen Studien wurden auf vorher bestimmte Ein- und Ausschlusskriterien untersucht und ggf. ausgeschlossen. Die aufgenommenen Studien wurden von zwei Bewertern unabhängig voneinander einem Evidenzgrad zugeordnet. Bei Uneinigkeiten wurde durch Diskussion ein Konsens gefunden. Anschließend wurde eine Metaanalyse durchgeführt, die sowohl eine Gesamtanalyse als auch mehrere Subgruppenanalysen beinhaltete. Die diagnostischen Richtigkeiten der einzelnen Studien wurden hierbei in Form der bivariaten Analyse ausgewertet, worin der zweidimensionale Effekt von Sensitivität und Spezifität Berücksichtigung fand.

Insgesamt wurden 46 Studien in die Analyse aufgenommen, davon benutzten 8 Studien die CT als diagnostischen Test, 10 Studien verwendeten den Ultraschall, 11 Studien das MRT und 4 Studien den AFP-Wert. 13 der Studien verwendeten mehr als

einen diagnostischen Test. Keiner der Studien konnte ein höherer Evidenzgrad als 2B zugeordnet werden. Häufige Gründe für die Herabstufung der Qualität der Studien waren fehlende Angaben über die Unabhängigkeit des Referenzstandards sowie über den Umgang mit nicht eindeutigen Befunden. Für die Ultraschalluntersuchung (7927 Patienten) konnte in der bivariaten Analyse ein Sensitivitätswert von 66,2% (KI 54,0% - 76,5%) bei einer zugehörigen Spezifität von 95,7% (KI 91,1% - 98,0%) gefunden werden. Für die klinische Praxis ist die kostengünstige und schnelle Untersuchung mit US bei befriedigender diagnostischer Richtigkeit somit v.a. für die Screeninguntersuchung geeignet. Die Analyse der Computertomographie (1578 Patienten) ergab für die Sensitivität 72,2% (KI 63,5% - 79,5%) und für die Spezifität 90,3% (KI 84,8% - 93,9%). In der Subgruppenanalyse zeigte sich kein Vorteil hinsichtlich Studien mit höherem Evidenzgrad oder Impaktfaktor, es zeigten sich aber auch hier gute Ergebnisse in den einzelnen Untergruppen. Für die Magnetresonanztomographie (882 Patienten) wurden in der bivariaten Analyse die besten Ergebnisse gefunden: Sensitivität 82% (KI 71,9% - 89%), Spezifität 90,3% (KI 83,5% - 94,4%). Bei einer mit der CT vergleichbaren Spezifität fand sich eine deutlich höhere Sensitivität. Die Subgruppenanalyse ergab gleich gute Ergebnisse in allen Untergruppen, jedoch fand sich auch hier keine Abweichung vom Gesamtergebnis. Die bivariate Analyse für die AFP-Diagnostik (2154 Patienten) ergab für die Sensitivität bei einem Cut-off >200 ng/ml einen Wert von 39,6% (KI 23,9% - 57,8%) und für die Spezifität einen Wert von 98,9% (KI 97,9% - 99,4%). Als sehr unspezifischer Marker bestätigte sich somit der relativ geringe Nutzen in der klinischen Praxis für die Diagnosestellung des HCC.

Für die Diagnostik des hepatozellulären Karzinoms zeigen sich v.a. für die CT- und MRT-Untersuchungen gute diagnostische Richtigkeiten. Die Kosten- und Zeitintensivität dieser Untersuchungen, sowie die Strahlenbelastung im CT stellen allerdings Limitationen in der täglichen Anwendung dar. Für die Screeninguntersuchung von Risikopatienten gilt somit weiterhin der Ultraschall als wichtige Voraussetzung. Dieser kann in Kombination mit der Erhebung des AFP-Werts ebenfalls gute Ergebnisse erzielen. Für die genaue Vergleichbarkeit der unterschiedlichen Diagnoseverfahren bedarf es aber auch in Zukunft der vermehrten Durchführung qualitativ hochwertiger Diagnosestudien.

## 7. ABSTRACT

Liver cancer is one of the ten most common causes of cancer death worldwide with an estimated 600.000 deaths each year [59]. Hepatocellular carcinoma as a histological subtype represents 85% of the liver cancers and thereby accounted for up to 5% of the worldwide cancer burden in the year 1990 [12]. Because of its 5 year survival rate of 6.5%, HCC is associated with a very poor prognosis [25]. In the last few years the incidence was rising constantly, especially in countries with a low prevalence of HCC like the USA or Western Europe. However, in parts of Asia or Africa where the prevalence of HCC is high, there was a decrease in the number of new cases of HCC [27]. It occurs almost twice as often in men than in women [40]. The most important risk factor for the development of HCC is the cirrhosis of the liver. A particularly high incidence of HCC occurs in patients with cirrhosis who also suffer from hepatitis C infection [28]. An early diagnosis is the decisive factor for the patients given the few therapeutic options.

The aim of this dissertation was to outline the currently available data for the diagnosis of HCC, to evaluate the existing studies concerning their methodology and quality and to determine their level of evidence in order to quantitatively summarize the results in a meta-analysis. Therefore a systematic literature research was conducted on PubMed, the thereby obtained studies were compared to predetermined inclusion and exclusion criteria and were excluded if necessary. The eligible studies were assigned to an evidence level by two independent reviewers. Any disagreement resulted in a consensus through discussion. Subsequently, a meta-analysis was conducted which included an overall analysis as well as several subgroup analyses. The diagnostic accuracies of each study were assessed by using the bivariate model to account for the two-dimensional effect of sensitivity and specificity.

In total, 46 studies were included, of these 8 studies used CT as a diagnostic test, 10 used ultrasound, 11 used MRI and 4 studies made use of AFP. 13 studies used more than one diagnostic test. None of the studies were assigned to a higher evidence level than 2B. In many cases the reasons for downgrading in quality were the lack of information about the independence of the reference standard as well as about the handling of inconclusive test results.

The bivariate analysis of ultrasound (7927 patients) resulted in a sensitivity of 66.2% (CI 54.0% - 76.5%) with a related specificity of 95.7% (CI 91.1% - 98.0%). Hence, the cost-efficient and fast method of US provides a satisfying diagnostic accuracy and is therefore convenient as a screening method in everyday clinical practice.

For the analysis of computed tomography (1578 patients) the analysis showed a sensitivity of 72.2% (CI 63.5% - 79.5%) and a specificity of 90.3% (CI 84.8% - 93.9%). The subgroup analysis did not display any advantages for studies with a better level of evidence or a higher impact factor, but it also showed good results for each subgroup.

The best results in the bivariate analysis were found for the magnetic resonance imaging (882 patients): sensitivity 82% (CI 71.9% - 89%), specificity 90.3% (CI 83.5% - 94.4%). With a specificity value similar to the one of the CT examination, MRI shows a considerably higher sensitivity value. The subgroup analysis revealed equally good results in each subgroup, however there were no divergence from the overall result.

The bivariate analysis of the AFP diagnosis (2154 patients) with a cut-off value of >200 ng/ml resulted in a sensitivity of 39.6% (CI 23.9% - 57.8%) and a related specificity of 98.9% (CI 97.9% - 99.4%). As a very unspecific diagnostic marker the relatively poor benefit of AFP for the clinical practice to diagnose HCC was confirmed.

For the diagnosis of HCC, good diagnostic accuracies are found particularly for CT and MRI examinations. Their advantages are limited by these procedures being costly and time-consuming, furthermore there is a significant radiation exposure to be considered for the CT examination. For the screening of patients at risk of developing HCC, the ultrasound imaging therefore still plays an important role. Its significance can be enhanced by combining it with the measurement of the biomarker AFP. For the exact comparison of the different diagnostic imaging techniques, well conducted diagnostic studies of high quality will be needed in the future.



## 8. ANHANG

### 8.1 Anlage 1: Tabelle 16 - Impaktfaktoren und Evidenzlevel der Studien

Name	Zeitschrift	Jahr	Impaktfaktor	Evidenzlevel
Maringhini	Cancer	1984	2,595	3B
Okazaki	Clin Oncol	1984	2,190*	4
Kobayashi	Hepatology	1985	4,994	4
Tanaka	Cancer	1986	2,330	3B
Dodd	Am J Roentgenol	1992	1,841	2B
Lopez	Int J Biol Marker	1996	1,288*	3B
Saada	Gut	1997	4,546	3B
Born	Fortschr Roentgen (Rofo– Fortschr Rontg)	1998	0,901	3B
Chalasani	Am J Gastroenterol	1999	2,941	3B
Gambarin-Gelwan	Am J Gastroenterol	2000	2,834	4
Lim	Am J Roentgenol	2000	1,863	2B
Peterson	Radiology	2000	4,13	2B
Kim CK	J Ultrasound Med	2001	1,258	4
Kim JE	Am J Roentgenol	2001	2,775	3B
Krinsky	Radiology	2001	4,759	2B
Mortèle	Eur Radiol	2001	1,321	3B
Rode	J Comput Assist Tomo	2001	1,302	3B
Tong	J Gastroenterol Hepatol	2001	1,258	2B
Trevisani	J Hepatol	2001	0,741	3B
Zacherl	Ann Surg	2001	6,674	2B
Bennett	Am J Roentgenol	2002	1,998	3B
Bhartia	Am J Roentgenol	2002	2,424	3B
De Lédinghen	Eur J Gastroenterol Hepatol	2002	1,337	2B
Marrero	Hepatology	2002	9,825	3B
Mori	Am J Roentgenol	2002	2,424	3B
Nguyen	Hepatology	2002	9,825	3B
Burrel	Hepatology	2003	9,503	2B
Libbrecht	Liver Transpl	2003	4,242	3B

<b>Name</b>	<b>Zeitschrift</b>	<b>Jahr</b>	<b>Impaktfaktor</b>	<b>Evidenzlevel</b>
Rickes	Eur J Gastroen Hepat	2003	1,578	2B
Teefey	Radiology	2003	4,815	3B
Valls	Am J Roentgenol	2004	2,380	3B
Lauenstein	Am J Roentgenol	2007	2,470	2B
Xu	Journal of Clinical Ultrasound	2007	0,645	2B
Choi	Am J Roentgenol	2008	2,940	3B
Compagnon	Transplantation	2008	3,816	3B
Forner	Hepatology	2008	11, 355	2B
Jang	Eur J Radiol	2009	2,645	3B
Yoo	Invest Radiol	2009	4,850	4
Luca	Eur Radiol	2010	3,594	3B
Talbot	J Nucl Med	2010	7,022	3B
Addley	Clin Radiol	2011	1,952	2B
Becker-Weidmann	Radiology	2011	5,726	3B
Haradome	J Magn Reson Image	2011	2,698	2B
Lee	Am J Roentgenol	2011	2,775	3B
Sano	Radiology	2011	5,726	2B
Suh	Am J Roentgenol	2011	2,775	4
Yu	Clin Gastroenterol Hepatol	2011	5,627	2B

\*Angegeben ist der 5-year Impact Factor

## 8.2 Anlage 2: Fragebogen zur Ermittlung des Evidenzgrades

<b>EBM Erhebungsbogen für diagnostische Studien</b>		Referenz-Nummer:	<div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black; margin: 0 5px;"></div> <div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black; margin: 0 5px;"></div> <div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black; margin: 0 5px;"></div> <div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black;"></div>
Erstautor:		Evidenzgrad:	<div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black;"></div>
Publikationsjahr:	<div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black; margin: 0 5px;"></div> <div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black; margin: 0 5px;"></div> <div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black; margin: 0 5px;"></div> <div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black;"></div>		
Land:		EBM ermittelt durch (Initialen):	<div style="display: inline-block; width: 20px; height: 20px; border: 1px solid black;"></div>

	Ja	Nein
Wurde ein prospektives Design verwendet?		
Wurde die Studienpopulation durch Einschluss-/Ausschlusskriterien eindeutig definiert?		
War das Studienkollektiv repräsentativ?		
Wurde bei allen Patienten ein Referenzstandard durchgeführt?		
War der Referenzstandard unabhängig? (Hatten die Beurteiler des Referenzstandards keine Kenntnis der Ergebnisse des diagnostischen Tests?)		
Waren die Beurteiler des diagnostischen Tests verblindet gegenüber dem Referenzstandard?		
Wurde der diagnostische Test unabhängig von weiteren klinischen Informationen (Anamnese, körperliche Untersuchung, Laborergebnisse) bewertet?		
Wurden klare Kriterien für die Auswertung des Referenzstandards definiert? (z.B. pathologische Kriterien für die Entscheidung Appendizitis ja/nein?)		
War das Follow-up gut?		
Wurde der Umgang mit nicht eindeutigen Befunden beschrieben?		

**Bemerkungen:**

### **8.3 Anlage 3: Tabelle 17 – Ausschlussgründe für gefundene Studien**

<b>Name</b>	<b>Jahr</b>	<b>Land</b>	<b>Ausschlussgrund</b>
Akai, H.	2011	Japan	Keine Angabe von Spezifität
Amano, S.	2003	Japan	Keine HCC-Diagnostik
Baek, C.K.	2012	Südkorea	Sequentieller Test
Bargellini, I.	2011	Italien	Review
Bartolozzi, C.	2011	Italien	Narrative Overview
Beadsmoore, C.J.	2005	UK	Fallbericht
Bhattacharjya, S.	2004	UK	Sequentieller Test
Bialecki, E.S.	2006	USA	Keine Histopathologie als Referenzstandard
Bolondi, L.	2005	Italien	Keine HCC-Diagnostik
Brancatelli, G.	2001	USA/Italien	Keine HCC-Diagnostik
Caturelli, E.	1996/2002/2004	Italien	Keine Histopathologie als Referenzstandard
Ch Yu, S.	1997	China	Diagnostischer Test unpassend (Biopsie)
Chen, J.H.	1999	Taiwan	Kein Referenzstandard
Choi, B.I.	2011	Südkorea	Narrative Overview
Chou, C.T.	2010	Taiwan	Keine HCC-Diagnostik
Chou, C.T.	2011	Taiwan	Keine HCC-Diagnostik
Colombo, M.	1991	Italien	Zusammengesetzter Referenzstandard
Compagnon, P.	2008	Frankreich	Angabe von Sensitivität u. Spezifität nicht nachvollziehbar
Cwikla, J.B.	2001	England	Keine HCC-Diagnostik
D’Onofrio, M.	2004	Italien	Keine HCC-Diagnostik
Di Martino, M.	2010	Italien	Keine Histopathologie als Referenzstandard
Durand, F.	2001	Frankreich	Diagnostischer Test unpassend (Biopsie)
Edeline, J.	2012	Frankreich	Keine Histopathologie als Referenzstandard
Elias, D.	2005	Frankreich	Unpassendes Patientenkollektiv

Enomoto, S.	2011	Japan	Diagnostischer Test diente Differenzierung des Tumors, nicht der Entdeckung
Forner, A.	2006	USA	Narrative Overview
Frieser, M.	2011	Deutschland	Beurteilung des Behandlungserfolgs
Fukuda, H.	2011	Japan	Keine HCC-Diagnostik
Furlan, A.	2011	USA	Keine Histopathologie als Referenzstandard
Golfieri, R.	2009	Italien	Unpassendes Patientenkollektiv
Golfieri, R.	2011	Italien	Sequentieller Test
Goshima, S.	2006/2011	Japan	Keine Histopathologie als Referenzstandard
Griffits, J.	2010	USA	Tierversuch
Han, E.	2010	Taiwan	Laborexperiment
Hayashi, P.H.	2004	USA	Uneinheitlicher diagnostischer Test
He, Y.M.	2005	China	Diagnostischer Test unpassend (Biopsie)
Herszenyi, L.	1995	Italien	Keine Histopathologie als Referenzstandard; Diagnostischer Test unpassend (Biopsie)
Holland, A.E.	2005	USA	Keine Angabe von Sens./Spez.
Huang, G.T.	1996	Taiwan	Diagnostischer Test unpassend (Biopsie)
Huo, T.I.	2004	Taiwan	Keine ausreichendes Follow-up
Imai, Y.	2008	Japan	Keine Angabe von Spezifität
Jiang, Y.	2010	China	Chinesischer Originaltext
Kawada, M.	2010	Japan	Keine Angabe von Spezifität
Kim, H.C.	2008	Südkorea	Keine HCC-Diagnostik
Kim, J.I.	2008	Südkorea	Zusammengesetzter Referenzstandard
Kim, S.H.	2009	Südkorea	Sequentieller Test
Kim, T.	2001	USA	Keine HCC-Diagnostik
Kim, T.K.	2011	Kanada	Unpassender Referenzstandard
Kim, Y.K.	2005	Südkorea	Sequentieller Test
Kim, Y.K.	2008	Südkorea	Kein ausreichendes Follow-up

Kim, Y.K. (Clin Radiol)	2011	Südkorea	Kein ausreichendes Follow-up
Kim, Y.K. (Invest Radiol)	2011	Südkorea	Sequentieller Test
Kogita, S.	2010	Japan	Keine Angabe von Spezifität
Kumano, S.	2009	Japan	Keine Histopathologie als Referenzstandard
Kwak, H.S.	2004	Südkorea	Sequentieller Test
Lee, M.H.	2011	Südkorea	Diagnostischer Test nicht bei allen Patienten durchgeführt
Lee, S.	2010	Südkorea	Keine HCC-Diagnostik
Lencioni, R.	1996	Italien	Keine HCC-Diagnostik; Sequentieller Test
Leoni, S.	2010	Italien	Sequentieller Test
Lin, M.T.	2010	Taiwan	Keine Angabe von Spezifität
Liu, Q.	2010	China	Keine HCC-Diagnostik
Liu, Q.Y.	2010	China	Diagnostischer Test nicht bei allen Patienten durchgeführt
Long, X.	2009	China	Chinesischer Originaltext
Long, X.Y.	2010	China	Unpassendes Patientenkollektiv
Ma, Y.	2008	China	Chinesischer Originaltext
Macaron, C.	2010	USA	Modellannahme
Mandai, M.	2011	Japan	Zusammengesetzter Referenzstandard
Marrero, J.A.	2005	USA	Zusammengesetzter Referenzstandard
Martie, A.	2011	Rumänien	Keine Angabe von Spezifität
Maturen, K.E.	2006	USA	Keine HCC-Diagnostik
Mion, F.	1996	Frankreich	Keine HCC-Diagnostik
Mita, K.	2010	Japan	Keine Angabe von Spezifität
Mok, T.S.K.	2004	China	Sequentieller Test
Nikolau, C.	2004	Spanien	Zusammengesetzter Referenzstandard
Noguchi, Y.	2003	Japan	Zusammengesetzter Referenzstandard
Okada, M.	2010	Japan	Keine HCC-Diagnostik

Okamoto, D.	2012	Japan	Keine HCC-Diagnostik
Park, G.	2009	Südkorea	Kein ausreichendes Follow-up
Park, M.J.	2010	Südkorea	Outcome-Studie
Park, Y.	2010	Südkorea	Sequentieller Test
Rhee, H.	2012	Südkorea	Keine HCC-Diagnostik
Sangiovanni, A.	2010	Italien	Unpassende Ergebnisdarstellung
Sangiovanni, A.	2004	Italien	Prognose-Studie
Seitz, K.	2010	Deutschland	Zusammengesetzter Referenzstandard
Sersté, T.	2011	Frankreich	Kein ausreichendes Follow-up
Singh, P.	2007	USA	Sequentieller Test
Sorrentino, P.	2011	Italien	Keine Histopathologie als Referenz- standard
Sotiropoulos, G.C.	2008	Deutschland	Keine HCC-Diagnostik
Stoker, J.	2002	Niederlande	Referenzstandard unklar
Strobel, D.	2005	Deutschland	Zusammengesetzter Referenzstandard
Strobel, D.	2009	Deutschland	Zusammengesetzter Referenzstandard
Sugimachi, K.	2011	Japan	Studienprotokoll unklar
Sun, H.Y.	2010	Südkorea	Zusammengesetzter Referenzstandard
Tarhan, C.	2011	Türkei	Keine HCC-Diagnostik
Vilana, R.	2003	Spanien	Zusammengesetzter Referenzstandard
Von Herbay, A.	2009	Deutschland	Keine HCC-Diagnostik
Wang, C.S.	2006	China	Chinesischer Originaltext
Wang, J.H.	2006	Taiwan	Keine Histopathologie als Referenz- standard
Wang, P.	2008	China	Diagnostischer Test unpassend (Biopsie)
Wen, Y.L.	2004	China	Zusammengesetzter Referenzstandard
Xiao, J.D.	2010	China	Keine HCC-Diagnostik
Yasui, K.	2011	Japan	Keine HCC-Diagnostik

Yoon, J.H.	2010	Südkorea	Keine Histopathologie als Referenzstandard
Yu, J.S.	2011	Südkorea	Keine HCC-Diagnostik
Zhou, K.R.	2002	China	Diagnostischer Test nicht bei allen Patienten durchgeführt



## 8.4 Anlage 4: Tabelle 18 – Levels of Evidence [60]

### Oxford Centre for Evidence-based Medicine Levels of Evidence (March 2009)

(for definitions of terms used see glossary at <http://www.cebm.net/?o=1116>)

Level	Therapy/Prevention, Aetiology/Harm	Prognosis	Diagnosis	Differential diagnosis/symptom prevalence study	Economic and decision analyses
1a	SR (with homogeneity*) of RCTs	SR (with homogeneity*) of inception cohort studies; CDR <sup>†</sup> validated in different populations	SR (with homogeneity*) of Level 1 diagnostic studies; CDR <sup>†</sup> with 1b studies from different clinical centres	SR (with homogeneity*) of prospective cohort studies	SR (with homogeneity*) of Level 1 economic studies
1b	Individual RCT (with narrow Confidence Interval <sup>‡</sup> )	Individual inception cohort study with > 80% follow-up; CDR <sup>†</sup> validated in a single population	Validating** cohort study with good <sup>+++</sup> reference standards; or CDR <sup>†</sup> tested within one clinical centre	Prospective cohort study with good follow-up****	Analysis based on clinically sensible costs or alternatives; systematic review(s) of the evidence; and including multi-way sensitivity analyses
1c	All or none <sup>§</sup>	All or none case-series	Absolute SpPins and SnNouts <sup>††</sup>	All or none case-series	Absolute better-value or worse-value analyses ++++
2a	SR (with homogeneity*) of cohort studies	SR (with homogeneity*) of either retrospective cohort studies or untreated	SR (with homogeneity*) of Level >2 diagnostic studies	SR (with homogeneity*) of 2b and better studies	SR (with homogeneity*) of 2b Level >2 economic studies

		control groups in RCTs			
2b	Individual cohort study (including low quality RCT; e.g., <80% follow-up)	Retrospective cohort study or untreated control patients in an RCT; Derivation of CDR <sup>†</sup> validated on split-sample\$§§ only	Exploratory** of good+++ reference standards; CDR <sup>†</sup> after derivation, or validated only on split-sample\$§§ or databases	Retrospective cohort study, or poor follow-up	Analysis based on clinically sensible costs or alternatives; limited review(s) of the evidence, or single studies; and including multi-way sensitivity analyses
2c	"Outcomes" Research; Ecological studies	"Outcomes" Research		Ecological studies	Audit or outcomes research
3a	SR (with homogeneity*) of case-control studies		SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies	SR (with homogeneity*) of 3b and better studies
3b	Individual Case-Control Study		Non-consecutive study; or without consistently applied reference standards	Non-consecutive cohort study, or very limited population	Analysis based on limited alternatives or costs, poor quality estimates of data, but including sensitivity analyses incorporating clinically sensible variations.
4	Case-series (and poor quality cohort and case-control	Case-series (and poor quality prognostic	Case-control study, poor or independent	Case-series or superseded reference standards	Analysis with no sensitivity analysis

	studies§§)		cohort studies***)		reference standard					
5	Expert	opinion	Expert	opinion	Expert	opinion	Expert	opinion	Expert	opinion
	without	explicit	without	explicit	without	explicit	without	explicit	without	explicit
	critical appraisal, or	critical appraisal,	critical appraisal, or	critical appraisal,	critical appraisal, or	critical appraisal, or	critical appraisal, or	critical appraisal, or	critical appraisal,	critical appraisal,
	based on physiology, or	based on	based on	based on	based on	based on	based on	based on	based on	based on
	bench research or	physiology,	physiology, bench	physiology, bench	bench research or	bench research or	bench research or	bench research or	economic theory	economic theory
	"first principles"	bench research	research or "first	research or "first	"first principles"	"first principles"	"first principles"	"first principles"	or	"first
		or "first principles"	principles"	principles"					principles"	

Produced by Bob Phillips, Chris Ball, Dave Sackett, Doug Badenoch, Sharon Straus, Brian Haynes, Martin Dawes since November 1998. Updated by Jeremy Howick March 2009.

## 9. LITERATUR

- [1] – Al Hasani F, Knoepfli M, Gemperli A, Kollar A, Banz V, Kettenbach J, Jüni P, Dufour JF. Factors affecting screening for hepatocellular carcinoma. *Annals of Hepatology*. 2014; 13(2):204-210
- [2] – Andersen J, Belmont J, Cho CT. Journal impact factor in the era of expanding literature. *Journal of Microbiology, Immunology and Infection*. 2006;39:436-443
- [3] – Andreana L, Isgrò G, Pleguezuelo M, Germani G, Burroughs AK. Surveillance and diagnosis of hepatocellular carcinoma in patients with cirrhosis. *World Journal of Hepatology*. 2009;1(1):48-61
- [4] – Ayuso C, Rimola J, García-Criado A. Imaging of HCC. *Abdominal Imaging*. 2012;37:215-230
- [5] – Baig JA, Alam JM, Mahmood SR, Baig M, Shaheen R, Sultana I, Waheed A. Hepatocellular Carcinoma (HCC) and diagnostic significance of  $\alpha$ -fetoprotein (AFP). *Journal of Ayub Medical College*. 2009; 21(1):72-75
- [6] – Bargellini I. Hepatocellular carcinoma: MR staging and therapeutic decisions. *Abdominal Imaging*. 2012;37:231-238
- [7] – Bazarah SM, Peltekian KM, McAlister VC, Bitter-Suermann H, MacDonald AS. Utility of MELD and Child–Turcotte–Pugh scores and the Canadian waitlisting algorithm in predicting shortterm survival after liver transplant. *Clinical and Investigative Medicine*. 2004;27(4):162-167
- [8] – Befeler AS, Di Bisceglie AM. Hepatocellular carcinoma: Diagnosis and treatment. *Gastroenterology*. 2002;122:1609-1619
- [9] – Bialecki ES, Di Bisceglie AM. Diagnosis of Hepatocellular carcinoma. *HPB (Oxford)*. 2005;7(1):26-34
- [10] – Bolog N, Andreisek G, Oancea I, Mangrau A. CT and MR Imaging of Hepatocellular Carcinoma. *Journal of Gastrointestinal and Liver Disease*. 2011;20(2):181-189
- [11] – Bolondi L, Gaiani S, Celli N, Golfieri R, Grigioni WF, Leoni S, Venturi AM, Piscaglia F. Characterization of Small Nodules in Cirrhosis by Assessment of Vascularity: The Problem of Hypovascular Hepatocellular Carcinoma. *Hepatology*. 2005;42(1):27-34
- [12] – Bosch FX, Ribes J, Borràs J. Epidemiology of Primary Liver Cancer. *Seminars in liver disease*. 1999;19(3):271-285
- [13] – Bosetti C, Turati F, La Vecchia C. Hepatocellular carcinoma epidemiology. *Best Practice and Research Clinical Gastroenterology*. 2014; 28(5):753-770

- [14] – Bruix J, Sherman M. Management of Hepatocellular Carcinoma. *Hepatology*. 2005;42(5):1208-1236
- [15] – Bruix J, Sherman M. Management of Hepatocellular Carcinoma: An Update. *Hepatology*. 2011;53(3):1020-1022
- [16] – Burrel M, Reig M, Forner A, Barrufet M, de Lope CR, Tremosini S, Ayuso C, Llovet JM, Real MI, Bruix J. Survival of patients with hepatocellular carcinoma treated by transarterial chemoembolisation (TACE) using Drug Eluting Beads. Implications for clinical practice and trial design. *Journal of Hepatology*. 2012;56(6):1330-1335
- [17] – Cho JY, Paik YH, Sohn W, Cho HC, Gwak GY, Choi MS, Lee JH, Koh KC, Paik SW, Yoo BC. Patients with chronic hepatitis B treated with oral antiviral therapy retain a higher risk for HCC compared with patients with inactive stage disease. *Gut*. 2014;63(12):1943-1950
- [18] – Cintonino D, Spada M, Gruttadauria S, Riva S, Luca A, Volpes R, Vizzini G, Arcadipane A, Henderson K, Verzaro R, Scotti Foglieni C, Gridelli B. *Transplantation Proceedings*. 2006;38(4):1096-1098
- [19] – Colombo M, Iavarone M. Role of antiviral treatment for HCC prevention. *Best Practice and Research Clinical Gastroenterology*. 2014;28(5):771-781
- [20] – Crocetti L, de Baere T, Lencioni R. Quality Improvement Guidelines for Radiofrequency Ablation of Liver Tumors. *CardioVascular and Interventional Radiology*. 2010;33:11-17
- [21] – Cuchetti A, Trevisani F, Pecorelli A, Erroi V, Farinati F, Ciccarese F, Rapaccini GL, Marco MD, Caturelli E, Giannini EG, Zoli M, Borzio F, Cabibbo G, Felder M, Gasbarrini A, Sacco R, Foschi FG, Missale G, Morisco F, Baroni GS, Virdone R, Bernardi M, Pinna AD, Italian Liver Cancer (ITA.LI.CA) group. Estimation of lead-time bias and its impact on the outcome of surveillance for the early diagnosis of hepatocellular carcinoma. *Journal of Hepatology*. 2014;61(2):333-341
- [22] – D'Alessandro LA, Meyer R, Klingmüller U. Hepatocellular carcinoma: a systems biology perspective. *Frontiers in Physiology*. 2013;4:1-6
- [23] – De Lemos AS, Chung RT. Hepatitis C treatment: an incipient therapeutic revolution. *Trends in Molecular Medicine*. 2014;20(6):315-321
- [24] – De Queiroz MR, Neto MJ, Garcia RG, Rahal Junior A, Salvalaggio P, Funari MB. High correlation between microbubble contrast-enhanced ultrasound, magnetic resonance and histopathology in the evaluation of hepatocellular carcinoma. *Einstein (Sao Paulo)*. 2013; 11(4):500-506
- [25] – Denzer U, Guthoff A, Wege H, Hillert C, Krupski G. Leberkrankheiten Teil II - Das hepatozelluläre Karzinom. *Hamburger Ärzteblatt*. 2006;4:207-213
- [26] – De Vos Irvine H, Goldberg D, Hole DJ, McMenamin J. Trends in primary liver cancer. *Lancet*. 1998;351(9097):215-216

- [27] – El-Serag HB, Rudolph KL. Hepatocellular Carcinoma: Epidemiology and Molecular Carcinogenesis. *Gastroenterology*. 2007;132:2557–2576
- [28] – Fattovich G, Stroffolini T, Zagni I, Donato F. Hepatocellular Carcinoma in Cirrhosis: Incidence and Risk Factors. *Gastroenterology*. 2004;127:S35-S50
- [29] – Floriani I, D’Onofrio M, Rulli E, Chen M.-H., Li R, Musicco L. Performance of Imaging Modalities in the Diagnosis of Hepatocellular Carcinoma: a Systematic Review and Meta-Analysis. *Ultraschall in der Medizin*. 2013; 34(5):454-462
- [30] – Fonseca AL, Cha CH. Hepatocellular Carcinoma: A Comprehensive Overview of Surgical Therapy. *Journal of Surgical Oncology*. 2014;110(6):712-719
- [31] – Forner A, Hessheimer AJ, Real MI, Bruix J. Treatment of hepatocellular carcinoma. *Critical Reviews in Oncology/Hematology*. 2006; 60:89-98
- [32] – Garfield E. How can impact factors be improved? *British Medical Journal*. 1996;313:411-413
- [33] – Garfield E. Journal impact factor: a brief review. *Canadian Medical Association Journal*. 1999;161(8):979-980
- [34] – Garfield E. The History and Meaning of the Journal Impact Factor. *Journal of the American Medical Association*. 2006; 295(1):90-93
- [35] – Greten T.F., Malek N.P., Schmidt S., Arends J., Bartenstein P., Bechstein W., Bernatik T., Bitzer M., Chavan A., Dollinger M., Domagk D., Drognitz O., Dux M., Farkas S., Folprecht G., Galle P., Geißler M., Gerken G., Habermehl D., Helmberger T., Herfarth K., Hoffmann R.-T., Holtmann M., Huppert P., Jakobs T., Keller M., Klempnauer J., Kolligs F., Körber J., Lang H., Lehner F., Lordick F., Lubienski A., Manns M.P., Mahnken A., Möhler M., Mönch C., Neuhaus P., Niederau C., Ocker M., Otto G., Pereira P., Pott G., Riemer J., Ringe K., Ritterbusch U., Rummeny E., Schirmacher P., Schlitt H.J., Schlottmann K., Schmitz V., Schuler A., Schulze-Bergkamen H., Schweinitz von D., Seehofer D., Sitter H., Straßburg C.P., Stroszczyński C., Strobel D., Tannapfel A., Trojan J., van Thiel I., Vogel A., Wacker F., Wedemeyer H., Wege H., Weinmann A., Wittekind C. Wörmann B., Zech C. Diagnostik und Therapie des Hepatozellulären Krazinoms, Langversion 1.0, AWMF Registrierungsnummer: 032-0530L, <http://leitlinienprogramm-onkologie.de/Leitlinien.7.0.html> (Stand: 11.09.2013)
- [36] – Guang Y, Xie LM, Ding H, Cai AL, Huang Y. Diagnosis value of focal liver lesions with SonoVue®-enhanced ultrasound compared with contrast-enhanced computed tomography and contrast-enhanced MRI: a meta-analysis. *Journal of Cancer Research and Clinical Oncology*. 2011;137(11):1595-1605
- [37] – Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne J. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics*. 2007;8(2):239-251
- [38] – Heverhagen JT, Pfestroff K, Heverhagen AE, Klose KJ, Kessler K, Sitter H. Diagnostic accuracy of magnetic resonance imaging: a prospective evaluation of

patients with suspected appendicitis (diamond). *Journal of Magnetic Resonance Imaging*. 2012;35(3):617-623

[39] – Hsu CY, Hsia CY, Huang YH, Su CW, Lin HC, Lee PC, Loong CC, Chiang JH, Huo TI, Lee SD. Selecting an Optimal Staging System for Hepatocellular Carcinoma: comparison of 5 currently used prognostic models. *Cancer*. 2012; 116(12):3006–3014

[40] – Jemal A, Bray F, Center M, Ferlay J, Ward E, Forman D. Global cancer statistics. *Cancer Journal for Clinicians*. 2011;61(2):69-90

[41] – Jin YJ, Lee JW, Park SW, Lee JI, Lee DH, Kim YS, Cho SG, Jeon YS, Lee KY, Ahn SI. Survival outcome of patients with spontaneously ruptured hepatocellular carcinoma treated surgically or by transarterial embolization. *World Journal of Gastroenterology*. 2013;19(28):4537-4544

[42] – Kim H, Park YN. Role of biopsy sampling for diagnosis of early and progressed hepatocellular carcinoma. *Best Practice and Research Clinical Gastroenterology*. 2014;28(5):813-829

[43] – Kudo M. Adjuvant Therapy after Curative Treatment for Hepatocellular Carcinoma. *Oncology*. 2011;81(Suppl.1):50-55

[44] – Kudo M, Han G, Finn RS, Poon RT, Blanc JF, Yan L, Yang J, Lu L, Tak WY, Yu X, Lee JH, Lin SM, Wu C, Tanwandee T, Shao G, Walters IB, Dela Cruz C, Poulart V, Wang JH. Brivanib as Adjuvant Therapy to Transarterial Chemoembolization in Patients With Hepatocellular Carcinoma: A Randomized Phase III Trial. *Hepatology*. 2014;60(5):1697-1707

[45] – Kumar A, Acharya SK, Singh SP, Saraswat VA, Arora A, Duseja A, Goenka MK, Jain D, Kar P, Kumar M, Kumaran V, Mohandas KM, Panda D, Paul SB, Ramachandran J, Ramesh H, Rao PN, Shah SR, Sharma H, Thandassery RB. The Indian National Association for Study of the Liver (INASL) Consensus on Prevention, Diagnosis and Management of Hepatocellular Carcinoma in India: The Puri Recommendations. *Journal of Clinical and Experimental Hepatology*. 2014;3(Suppl.3):S1-S24

[46] – Lai EC, Lau WY. Spontaneous rupture of hepatocellular carcinoma: a systematic review. *Archives of Surgery*. 2006;141(2):191-198

[47] – Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia, Critical Care & Pain*. 2008;8(6):221-223

[48] – Lencioni R, Piscaglia F, Bolondi L. Contrast-enhanced ultrasound in the diagnosis of hepatocellular carcinoma. *Journal of Hepatology*. 2008;48:848-857

[49] – Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH, Bossuyt PM. Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests. *Journal of the American Medical Association*. 1999; 282(11):1061-1066

- [50] – Lin HM, Lei LM, Zhu J, Li GL, Min J. Risk factor analysis of perioperative mortality after ruptured bleeding in hepatocellular carcinoma. *World Journal of Gastroenterology*. 2014; 20(40):14921-14926
- [51] – Llovet JM, Ricci S, Mazzaferro v, Hilgard P, Gane E, Blanc JF, de Oliveira AC, Santoro A, Raoul JL, Forner A, Schartz M, Porta C, Zeuzem S, Bolondi L, Greten TF, Galle PR, Seitz JF, Borbath I, Häussinger D, Giannaris T, Shan M, Moscovici M, Voliotis D, Bruix J. Sorafenib in advanced Hepatocellular Carcinoma. *The New England Journal of Medicine*. 2008;359(4):378-390
- [52] – Lok AS, Gardiner DF, Lawitz E, Martorell C, Everson GT, Ghalib R, Reindollar R, Rustgi V, McPhee F, Wind-Rotolo M, Persson A, Zhu K, Dimitrova DI, Eley T, Guo T, Grasela DM, Pasquinelli C. Preliminary study of two antiviral agents for Hepatitis C Genotype 1. *The New England Journal of Medicine*. 2012.366(3):216-224
- [53] – Mazzaferro V, Llovet JM, Miceli R, Bhoori S, Schiavo M, Mariani L, Camerini T, Roayaie S, Schwartz ME, Grazi GL, Adam R, Neuhaus P, Salizzoni M, Bruix J, Forner A, De Carlis L, Cillo U, Burroughs AK, Troisi R, Rossi M, Gerunda GE, Lerut J, Belghiti J, Boin I, Gugenheim J, Rochling F, Van Hoek B, Majno P. Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. *The Lancet Oncology*. 2009;10(1):35-43
- [54] – McGlynn KA, Tsao L, Hsing AW, Devesa SS, Fraumeni JF Jr. International trends and patterns of primary liver cancer. *International Journal of Cancer*. 2001;94(2):290-296
- [55] – McGlynn KA, London WT. Epidemiology and natural history of hepatocellular carcinoma. *Best Practice and Research Clinical Gastroenterology*. 2005;19(1):3-23
- [56] – Murata S, Mine T, Sugihara F, Yasui D, Yamaguchi H, Ueda T, Onozawa S, Kumita SI. Interventional treatment for unresectable hepatocellular carcinoma. *World Journal of Gastroenterology*. 2014;20(37): 13453-13465
- [57] – Opthof T. Sense and nonsense about the impact factor. *Cardiovascular Research*. 1997;33:1-7
- [58] – Papatheodoridis GV, Lampertico P, Manolakopoulos S, Lok A. Incidence of hepatocellular carcinoma in chronic hepatitis B patients receiving nucleos(t)ide therapy: A systematic review. *Journal of Hepatology*. 2010;53(2):348-356
- [59] – Perz JF, Armstrong GL, Farrington LA, Hutin YJF, Bell BP. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *Journal of Hepatology*. 2006; 45:529-538
- [60] – Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, Dawes M, Howick J.: Oxford Centre for Evidence-based Medicine Levels of Evidence (March 2009). URL: <http://www.cebm.net/?o=1025> (Stand:17.09.2013)



- [61] – Pineau P, Marchio A, Nagamori S, Seki S, Tiollais P, Dejean A. Homozygous Deletion Scanning in Hepatobiliary Tumor Cell Lines Reveals Alternative Pathways for Liver Carcinogenesis. *Hepatology*. 2003;37(4):852-861
- [62] – Reitsma JB, Glas AS, Rutjes A, Scholten R, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology*. 2005;58:982-990
- [63] – Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *Journal of Clinical Epidemiology*. 2009;62(8):797-806
- [64] – Ronot M, Vilgrain V. Hepatocellular carcinoma: Diagnostic criteria by imaging techniques. *Best Practice and Research Clinical Gastroenterology*. 2014;28(5):795-812
- [65] – Sacco R, Gadaleta-Caldarola G, Galati G, Lombardi G, Mazza GC, Cabibbo G. European Association for the Study of the Liver Hepatocellular Carcinoma summit 2014: old questions, new (or few) answers?. *Future Oncology*. 2014;10(10):1719-1721
- [66] – Saha S, Saint S, Christakis DA. Impact factor: a valid measure of journal quality? *Journal of the Medical Library Association*. 2003;91(1):42-46
- [67] – Schiemann U, Szücs-Farkas Z, Tshering D, Kunz A, Patak MA. Kontrastmittelultraschall (CEUS) bei Leberläsionen. Hand in Hand mit CT und MRI. *Schweizerisches Medizin-Forum*. 2009;9(50):915-917
- [68] – Schmidt RL, Factor RE. Understanding Sources of Bias in Diagnostic Accuracy Studies. *Archives of Pathology&Laboratory Medicine*. 2013;137:558-565
- [69] – Spangenberg HC, Zuber-Jerger I, Mohr L, Thimme R, Blum HE. Diagnostik und Therapie des hepatozellulären Karzinoms. *Deutsche Medizinische Wochenschrift*. 2004;129(8):368-372
- [70] – Shao W, Zhang F, Cong N, Li J, Song J. Transarterial chemoembolization combined with sorafenib for advanced hepatocellular carcinoma. *Oncology Letters*. 2014;8(5):2263-2266
- [71] – Sherman M. Surveillance for hepatocellular carcinoma. *Best Practice and Research Clinical Gastroenterology*. 2014;28(5):783-793
- [72] – Thorgeirsson SS. Hunting for tumor suppressor genes in liver cancer. *Hepatology*. 2003;37(4):739-741
- [73] – Thun, MJ. The global burden of cancer: priorities for prevention. *Carcinogenesis*. 2010;31(1):100-110
- [74] – Torzilli G, Belghiti J, Makuuchi M. Differences and Similarities in the Approach to Hepatocellular Carcinoma Between Eastern and Western Institutions. *Liver Transplantation*. 2004; 10(2 Suppl 1):S1-2

- [75] – Ulahannan SV, Duffy AG, McNeel TS, Kish JK, Dickie LA, Rahma OE, McGlynn KA, Greten TF, Altekruse SF. Earlier Presentation and Application of Curative Treatments in Hepatocellular Carcinoma. *Hepatology*. 2014;60(5):1637-1644
- [76] – Van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine*. 2002;21:589-624
- [77] – Wong GL, Chan HL, Mak CW, Lee SK, Ip ZM, Lam AT, Iu HW, Leung JM, Lai JW, Lo AO, Chan HY, Wong VW. Entecavir Treatment Reduces Hepatic Events and Deaths in Chronic Hepatitis B Patients With Liver Cirrhosis. *Hepatology*. 2013;58(5):1537-1547
- [78] – Walenbergh S, Koek GH, Bieghs V, Shiri-Sverdlov R. Non-alcoholic steatohepatitis: The role of oxidized low-density lipoproteins. *Journal of Hepatology*. 2013;58:801-810
- [79] – Weinstein S, Obuchowski NA, Lieber ML. Clinical Evaluation of Diagnostic Tests. *American Journal of Roentgenology*. 2005;184(1):14-19
- [80] – Willatt JM, Hussain HK, Adusumilli S, Marrero JA. MR imaging of Hepatocellular Carcinoma in the cirrhotic liver: Challenges and controversies. *Radiology*. 2008;247(2):311-330
- [81] – Willis BH. Spectrum bias—why clinicians need to be cautious when applying diagnostic test studies. *Family Practice*. 2008; 25:390-396
- [82] – Yoshimoto S, Loo TM, Atarashi K, Kanda H, Sato S, Oyadomari S, Iwakura Y, Oshima K, Morita H, Hattori M, Honda K, Ishikawa Y, Hara E, Ohtani N. Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature*. 2013; 499(7456):97-101
- [83] – Zidan A, Scheuerlein H, Schüle S, Settmacher Utz, Rauchfuss F. Epidemiological Pattern of Hepatitis B and Hepatitis C as Etiological Agents for Hepatocellular Carcinoma in Iran and Worldwide. *Hepatitis Monthly*. 2012;12(10HCC):e6894

## **9.2 In der Metaanalyse verwendete Studien**

- [M1] – Addley HC, Griffin N, Shaw AS, Mannelli L, Parker RA, Aitken S, Wood H, Davies S, Alexander GJ, Lomas DJ. Accuracy of hepatocellular carcinoma detection on multidetector CT in a transplant liver population with explants liver correlation. *Clinical Radiology*. 2011;66(4):349-356
- [M2] – Bayati N, Silverman AL, Gordon SC. Serum Alpha-Fetoprotein Levels and Liver Histology in Patients With Chronic Hepatitis C. *The American Journal of Gastroenterology*. 1998;93(12):2452-2456
- [M3] – Becker-Weidman DJ, Kalb B, Sharma P, Kitajima HD, Lurie CR, Chen Z, Spivey JR, Knechtle SJ, Hanish SI, Adsay NV, Farris AB 3<sup>rd</sup>, Martin DR. Hepatocellular Carcinoma Lesion Characterization: Single-Institution Clinical Performance Review of Multiphase

Gadolinium-enhanced MR Imaging – Comparison to Prior Same-Center Results after MR Systems Improvements. *Radiology*. 2011;261(3):824-833

[M4] – Bennett GL, Krinsky GA, Abitbol RJ, Kim SY, Theise ND, Teperman LW. Sonographic Detection of Hepatocellular Carcinoma and Dysplastic Nodules in Cirrhosis: Correlation of Pretransplantation Sonography and Liver Explant Pathology in 200 Patients. *American Journal of Roentgenology*. 2002;179(1):75-80

[M5] – Bhartia B, Ward J, Guthrie JA, Robinson PJ. Hepatocellular Carcinoma in Cirrhotic Livers: Double-Contrast Thin-Section MR Imaging with Pathologic Correlation of Explanted Tissue. *American Journal of Roentgenology*. 2003;180(3):577-584

[M6] – Born M, Layer G, Kreft B, Schwarz N, Schild H. MRT, CT und CTAP in der Diagnostik maligner Lebertumoren bei Leberzirrhose. *Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*. 1998;168(6):567-572

[M7] – Burrel M, Llovet JM, Ayuso C, Iglesias C, Sala M, Miquel R, Caralt T, Ayuso JR, Solé M, Sanchez M, Brú C, Bruix J. MRI Angiography Is Superior to Helical CT for Detection of HCC Prior to liver Transplantation: An Explant Correlation. *Hepatology*. 2003;38(4):1034-1042

[M8] – Chalasani N, Horlander JC, Said A, Hoen H, Kopecky KK, Stockberger SM, Manam R, Kwo PY, Lumeng L. Screening for Hepatocellular Carcinoma in Patients With Advanced Cirrhosis. *The American Journal of Gastroenterology*. 1999;94(10):2988-2993

[M9] – Choi SH, Lee JM, Yu NC, Suh KS, Jang JJ, Kim SH, Choi BI. Hepatocellular Carcinoma in Liver Transplantation Candidates: Detection With Gadobenate Dimeglumine-Enhanced MRI. *American Journal of Roentgenology*. 2008;191:529-536

[M10] – Colli A, Fraquelli M, Casazza G, Massironi S, Colucci A, Conte D, Duca P. Accuracy of Ultrasonography, Spiral CT, Magnetic Resonance, and Alpha-Fetoprotein in Diagnosing Hepatocellular Carcinoma: A Systematic Review. *American Journal of Gastroenterology*. 2006;101:513-523

[M11] – De Lédinghen V, Laharie D, Lecesne R, Le Bail B, Winnock M, Bernard PH, Saric J, Couzigou P, Balabaud C, Bioulac-Sage P, Drouillard J. Detection of nodules in liver cirrhosis: spiral computed tomography or magnetic resonance imaging? A prospective study of 88 nodules in 34 patients. *European Journal of Gastroenterology & Hepatology*. 2002;14(2):159-165

[M12] – Dodd GD, Miller WJ, Baron RL, Skolnick ML, Campbell WL. Detection of Malignant Tumors in End-Stage Cirrhotic Livers: Efficacy of Sonography as a Screening Technique. *American Journal of Roentgenology*. 1992;159:727-733

[M13]– Forner A, Vilana R, Ayuso C, Bianchi L, Solé M, Ayuso JR, Boix L, Sala M, Varela M, Llovet JM, Brú C, Bruix J. Diagnosis of Hepatic Nodules 20 mm or Smaller in Cirrhosis: Prospective Validation of the Noninvasive Diagnostic Criteria for Hepatocellular Carcinoma. *Hepatology*. 2008;47:97-104

- [M14] – Gambarin-Gelwan M, Wolf DC, Shapiro R, Schwartz ME, Min AD. Sensitivity of Commonly Available Screening Tests in Detecting Hepatocellular Carcinoma in Cirrhotic Patients Undergoing Liver Transplantation. *The American Journal of Gastroenterology*. 2000; 95(6):1535-1538
- [M15] – Haradome H, Grazioli L, Tinti R, Morone M, Motosugi U, Sano K, Ichikawa T, Kwee TC, Colagrande S. Additional Value of Gadoteric Acid-DTPA-Enhanced Hepatobiliary Phase MR Imaging in the Diagnosis of Early-Stage Hepatocellular Carcinoma: Comparison with Dynamic Triple-Phase Multidetector CT Imaging. *Journal of Magnetic Resonance Imaging*. 2011;34(1):69-78
- [M16] – Jang HJ, Kim TK, Wilson SR. Small nodules (1-2cm) in liver cirrhosis: Characterization with contrast-enhanced ultrasound. *European Journal of Radiology*. 2009;72:418-424
- [M17] – Kim CK, Lim JH, Lee WJ. Detection of Hepatocellular Carcinomas and Dysplastic Nodules in Cirrhotic Liver - Accuracy of Ultrasonography in Transplant Patients. *Journal of Ultrasound in Medicine*. 2001;20:99-104
- [M18] – Kim JE, Kim SH, Lee SJ, Rhim H. Hypervascular Hepatocellular Carcinoma 1 cm or Smaller in Patients With Chronic Liver Disease: Characterization With Gadoteric Acid-Enhanced MRI That Includes Diffusion-Weighted Imaging. *American Journal of Roentgenology*. 2011; 196(6):W758-765
- [M19] – Kobayashi K, Sugimoto T, Makino H, Kumagai M, Onoura M, Tanaka N, Kato Y, Hattori N. Screening Methods for Hepatocellular Early Detection of Carcinoma. *Hepatology*. 1985;5(6):1100-1105
- [M20] – Krinsky GA, Lee VS, Theise ND, Weinreb JC, Rofsky NM, Diflo T, Teperman LW. Hepatocellular Carcinoma and Dysplastic Nodules in Patients with Cirrhosis: Prospective Diagnosis with MR Imaging and Explantation Correlation. *Radiology*. 2001;219:445-454
- [M21] – Lauenstein TC, Salman K, Morreireira R, Heffron T, Spivey JR, Martinez E, Sharma P, Martin DR. Gadolinium-Enhanced MRI for Tumor Surveillance Before Liver Transplantation: Center-Based Experience. *American Journal of Roentgenology*. 2007;189(3):663-670
- [M22] – Lee MH, Kim SH, Park MJ, Park CK, Rhim H. Gadoteric Acid-Enhanced Hepatobiliary Phase MRI and High-b-Value Diffusion-Weighted Imaging to Distinguish Well-Differentiated Hepatocellular Carcinomas From Benign Nodules in Patients With Chronic Liver Disease. *American Journal of Roentgenology*. 2011; 197(5):W868-875
- [M23] – Libbrecht L, Bielen D, Verslype C, Vanbeckevoort D, Pirenne J, Nevens F, Desmet V, Roskams T. Focal Lesions in Cirrhotic Explant Livers: Pathological Evaluation and Accuracy of Pretransplantation Imaging Examinations. *Liver Transplantation*. 2002;8(9):749-761

- [M24] – Lim JH, Kim CK, Lee WJ, Park CK, Koh KC, Paik SW, Joh JW. Detection of Hepatocellular Carcinomas and Dysplastic Nodules in Cirrhotic Livers: Accuracy of Helical CT in Transplant Patients. *American Journal of Roentgenology*. 2000;175(3):693-698
- [M25] – Lopez JB, Balasegaram M, Thambyrajah V. Serum CA 125 as a marker of hepatocellular carcinoma. *The International Journal of Biological Markers*. 1996;11(3):178-182
- [M26] – Luca A, Caruso S, Milazzo M, Mamone G, Marrone G, Miraglia R, Maruzzelli L, Carollo V, Minervini MI, Vizzini G, Gruttadauria S, Gridelli B. Multidetector-row computed tomography (MDCT) for the diagnosis of hepatocellular carcinoma in cirrhotic candidates for liver transplantation: prevalence of radiological vascular patterns and histological correlation with liver explants. *European Radiology*. 2010; 20(4):898-907
- [M27] – Maringhini A, Cottone M, Sciarrino E, Marcenò MP, La Seta F, Rinaldi F, Pagliaro L. Ultrasonographic and Radionuclide Detection of Hepatocellular Carcinoma in Cirrhotics With Low Alpha - Fetoprotein Levels. *Cancer*. 1984;54:2924-2926
- [M28] – Marrero JA, Su GL, Wei W, Emick D, Conjeevaram HS, Fontana RJ, Lok AS. Des-Gamma Carboxyprothrombin Can Differentiate Hepatocellular Carcinoma From Nonmalignant Chronic Liver Disease in American Patients. *Hepatology*. 2003;37(5):1114-1121
- [M29] – Mori K, Scheidler J, Helmberger T, Holzknecht M, Schauer R, Schirren CA, Bittmann I, Dugas M, Reiser M. Detection of Malignant Hepatic Lesions Before Orthotopic Liver Transplantation: Accuracy of Ferumoxides-Enhanced MR Imaging. *American Journal of Roentgenology*. 2002;179:1045-1051
- [M30] – Mortelé KJ, De Keukeleire K, Praet M, Van Vlierberghe H, De Hemptinne B, Ros PR. Malignant focal hepatic lesions complicating underlying liver disease: dual-phase contrast-enhanced spiral CT sensitivity and specificity in orthotopic liver transplant patients. *European Radiology*. 2001;11(9):1631-1638
- [M31] – Nguyen MH, Garcia RT, Simpson PW, Wright TL, Keeffe EB. Racial Differences in Effectiveness of  $\alpha$ -Fetoprotein for Diagnosis of Hepatocellular Carcinoma in Hepatitis C Virus Cirrhosis. *Hepatology*. 2002;36(2):410-417
- [M32] – Okazaki N, Yoshida T, Yoshino M, Matue H. Screening of patients with chronic liver disease for hepatocellular carcinoma by ultrasonography. *Clinical Oncology*. 1984;10:241-246
- [M33] – Peterson MS, Baron RL, Marsh JW Jr, Oliver JH 3<sup>rd</sup>, Confer SR, Hunt LE. Pretransplantation Surveillance for Possible Hepatocellular Carcinoma in Patients with Cirrhosis: Epidemiology and CT-based Tumor Detection Rate in 430 Cases with Surgical Pathologic Correlation. *Radiology*. 2000;217:743-749

[M34] – Piantino P, Arrigoni A, Brunetto MR, Gindro T. Alpha-Fetoprotein in Hepatic Pathology and Hepatocarcinoma. *The Journal of Nuclear Medicine and allied Sciences*. 1989;33(Suppl.3):34-38

[M35] – Rickes S, Schulze S, Neye H, Ocran KW, Wermke W. Improved diagnosing of small hepatocellular carcinomas by echo-enhanced power Doppler sonography in patients with Cirrhosis. *European Journal of Gastroenterology & Hepatology*. 2003;15(8):893-900

[M36] – Rode A, Bancel B, Douek P, Chevallier M, Vilgrain V, Picaud G, Henry L, Berger F, Bizollon T, Gaudin JL, Ducerf C. Small Nodule Detection in Cirrhotic Livers: Evaluation with US, Spiral CT, and MRI and Correlation with Pathologic Examination of Explanted Liver. *Journal of Computer Assisted Tomography*. 2001;25(3):327-336

[M37] – Saada J, Bhattacharya S, Dhillon AP, Dick R, Burroughs AK, Rolles K, Davidson BR. Detection of small hepatocellular carcinomas in cirrhotic livers using iodised oil computed Tomography. *Gut*. 1997;41:404-407

[M38] – Sano K, Ichikawa T, Motosugi U, Sou H, Muhi AM, Matsuda M, Nakano M, Sakamoto M, Nakazawa T, Asakawa M, Fujii H, Kitamura T, Enomoto N, Araki T. Imaging Study of Early Hepatocellular Carcinoma: Usefulness of Gadoteric Acid-enhanced MR Imaging. *Radiology*. 2011;261(3):834-844

[M39] – Suh YJ, Kim MJ, Choi JY, Park YN, Park MS, Kim KW. Differentiation of Hepatic Hyperintense Lesions Seen on Gadoteric Acid-Enhanced Hepatobiliary Phase MRI. *American Journal of Roentgenology*. 2011;197(1):W44-52

[M40] – Talbot JN, Fartoux L, Balogova S, Nataf V, Kerrou K, Gutman F, Huchet V, Ancel D, Grange JD, Rosmorduc O. Detection of Hepatocellular Carcinoma with PET/CT: A Prospective Comparison of 18F-Fluorocholine and 18F-FDG in Patients with Cirrhosis or Chronic Liver Disease. *The Journal of Nuclear Medicine*. 2010;51(11):1699-1706

[M41] – Tanaka S, Kitamura T, Ohshima A, Umeda K, Okuda S, Ohtani T, Tatsuta M, Yamamoto K. Diagnostic Accuracy of Ultrasonography for Hepatocellular Carcinoma. *Cancer*. 1986;58(2):344-347

[M42]– Teefey S, Hildeboldt CC, Dehdashti F, Siegel BA, Peters MG, Heiken JP, Brown JJ, McFarland EG, Middleton WD, Balfe DM, Ritter JH. Detection of Primary Hepatic Malignancy in Liver Transplant Candidates: Prospective Comparison of CT, MR Imaging, US, and PET. *Radiology*. 2003;226(2):533-542

[M43] – Tong MJ, Blatt LM, Kao VW. Chronic Liver Diseases: Epidemiology, Pathophysiology, Diagnosis and Treatment - Surveillance for hepatocellular carcinoma in patients with chronic viral hepatitis in the United States of America. *Journal of Gastroenterology and Hepatology*. 2001;16:553-559

- [M44] – Trevisani F, D’Intino PA, Morselli-Labate AM, Mazzella G, Accogli E, Caraceni P, Domenicali M, De Notariis S, Roda E, Bernardi M. Serum  $\alpha$ -fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status. *Journal of Hepatology*. 2001;34:570-575
- [M45] – Valls C, Cos M, Figueras J, Andía E, Ramos E, Sánchez A, Serrano T, Torras J. Pretransplantation Diagnosis and Staging of Hepatocellular Carcinoma in Patients with Cirrhosis: Value of Dual-Phase Helical CT. *American Journal of Roentgenology*. 2004;182(4):1011-1017
- [M46] – Xu HX, Xie XY, Lu MD, Liu GJ, Xu ZF, Zheng YL, Liang JY, Chen LD. Contrast-Enhanced Sonography in the Diagnosis of Small Hepatocellular Carcinoma  $\leq 2$  cm. *Journal of Clinical Ultrasound*. 2008; 36(5):257-266
- [M47] – Yoo HJ, Lee JM, Lee JY, Kim SH, Kim SJ, Han JK, Choi BI. Additional Value of SPIO-Enhanced MR Imaging for the Noninvasive Imaging Diagnosis of Hepatocellular Carcinoma in Cirrhotic Liver. *Investigative Radiology*. 2009; 44(12):800-807
- [M48] – Yu NC, Chaudhari V, Raman SS, Lassman C, Tong MJ, Busuttil RW, Lu DS. CT and MRI Improve Detection of Hepatocellular Carcinoma, Compared With Ultrasound Alone, in Patients With Cirrhosis. *Clinical Gastroenterology and Hepatology*. 2011;9:161-167
- [M49] – Zacherl J, Pokieser P, Wrba F, Scheuba C, Prokesch R, Zacherl M, Längle F, Berlakovich GA, Mühlbacher F, Steininger R. Accuracy of Multiphasic Helical Computed Tomography and Intraoperative Sonography in Patients Undergoing Orthotopic Liver Transplantation for Hepatoma: What is the Truth?. *Annals of Surgery*. 2002;235(4):528-532

## 10. ABKÜRZUNGSVERZEICHNIS

<b>(S)ROC</b>	(Summary) Receiver Operating Characteristics
<b>AASLD</b>	American Association for the Study of Liver Diseases
<b>Abb.</b>	Abbildung
<b>AFP</b>	Alpha-Fetoprotein
<b>BCLC</b>	Barcelona Clinic for Liver Cancer
<b>Bzw.</b>	Beziehungsweise
<b>Ca.</b>	Circa
<b>CEUS</b>	Contrast enhanced Ultrasound
<b>CI</b>	Confidence Interval
<b>CLIP</b>	Cancer of the Liver Italian Program
<b>CP</b>	Child-Pugh
<b>CT</b>	Computertomographie
<b>CTAP</b>	CT Arterial Portography
<b>d.h.</b>	Das heißt
<b>DNA</b>	<i>Englisch für</i> Desoxyribonukleinsäure
<b>DTPA</b>	Diethylene Triamine Pentaacetic Acid
<b>DWI</b>	Diffusion weighted imaging
<b>EASL</b>	European Association for the Study of the Liver
<b>EBM</b>	Evidence Based Medicine
<b>EOB</b>	Ethoxybenzyl
<b>F-FDG</b>	<sup>18</sup> F-2-Fluordesoxyglucose
<b>FN</b>	Falsch negativ
<b>FP</b>	Falsch positiv
<b>ggf.</b>	Gegebenenfalls
<b>HBV</b>	Hepatitis-B-Virus
<b>HCC</b>	Hepatozelluläres Karzinom
<b>HCV</b>	Hepatitis-C-Virus
<b>IF</b>	Impaktfaktor
<b>INR</b>	International Normalized Ratio
<b>JIS</b>	Japanese Integrated Staging Score
<b>KI</b>	Konfidenzintervall
<b>KM</b>	Kontrastmittel
<b>MDCT</b>	Multi detector CT
<b>MeSH</b>	Medical Subject Heading



<b>mind.</b>	mindestens
<b>MRI</b>	Magnetic resonance imaging
<b>MRT</b>	Magnetresonanztomographie
<b>NASH</b>	<i>Englisch für</i> Nicht-alkoholische Steatohepatitis
<b>OP</b>	Operation
<b>PEI</b>	Perkutane Ethanolinjektion
<b>PET</b>	Positronen-Emissions-Tomographie
<b>PS</b>	Performance Status
<b>QUADAS</b>	Quality Assessment of Diagnostic Accuracy Studies
<b>RCT</b>	Randomized Controlled Trial
<b>RFA</b>	Radiofrequenzablation
<b>RN</b>	Richtig negativ
<b>RP</b>	Richtig positiv
<b>s.u.</b>	Siehe unten
<b>SDCT</b>	Single detector CT
<b>Sens</b>	Sensitivität
<b>sog.</b>	Sogenannt
<b>Spez</b>	Spezifität
<b>SPIO</b>	Superparamagnetic Iron Oxide
<b>SR</b>	Systematic Review
<b>TACE</b>	Transarterielle Chemoembolisation
<b>TNM</b>	Classification of Malignant Tumors
<b>u.</b>	und
<b>US</b>	Ultraschall
<b>USA</b>	United States of America
<b>v.a.</b>	Vor allem
<b>vs.</b>	versus
<b>WHO</b>	World Health Organization
<b>z.B.</b>	Zum Beispiel
<b>z.T.</b>	Zum Teil

## 11. ABBILDUNGSVERZEICHNIS

Abb. 1:	Das BCLC-Staging-System für das HCC mit daraus resultierender Therapieempfehlung	S. 18
Abb. 2:	Der Child-Pugh-Score	S. 20
Abb. 3:	Vierfeldertafel zur Bestimmung der Sensitivität und Spezifität	S. 31
Abb. 4:	Flow-Chart-Diagramm zum Suchvorgang der Studien der Basisliteratursuche	S. 37
Abb. 5:	Flow-Chart-Diagramm zum Suchvorgang der Ultraschall-Studien	S. 38
Abb. 6:	Flow-Chart-Diagramm zum Suchvorgang der CT-Studien	S. 39
Abb. 7:	Flow-Chart-Diagramm zum Suchvorgang der MRT-Studien	S. 40
Abb. 8:	Forest plots für Sensitivität und Spezifität der Ultraschall-Studien	S. 64
Abb. 9:	Forest plots für Sensitivität und Spezifität der CT-Studien	S. 66
Abb. 10:	Forest plots für Sensitivität und Spezifität der MRT-Studien	S. 68
Abb. 11:	Sensitivität und Spezifität von US, CT und MRT	S. 76

## 12. TABELLENVERZEICHNIS

Tabelle 1:	Einteilung der Evidenzgrade diagnostischer Studien	S. 28
Tabelle 2:	Daten der Studien mit US als diagnostischem Test	S. 48
Tabelle 3:	Daten der Studien mit CT als diagnostischem Test	S. 52
Tabelle 4:	Daten der Studien mit MRT als diagnostischem Test	S. 56
Tabelle 5:	Daten der Studien mit AFP als diagnostischem Test	S. 61
Tabelle 6:	Sensitivität und Spezifität für US	S.63
Tabelle 7:	Ergebnisse der Subgruppen-Analyse für US	S. 64
Tabelle 8:	Sensitivität und Spezifität für CT	S. 65
Tabelle 9:	Ergebnisse der Subgruppen-Analyse für CT	S. 65
Tabelle 10:	Sensitivität und Spezifität für MRT	S. 67
Tabelle 11:	Ergebnisse der Subgruppen-Analyse für MRT	S. 67
Tabelle 12:	Sensitivität und Spezifität für AFP	S. 69
Tabelle 13:	Analyse-Ergebnisse für die US-Studien in Bezug auf den Impaktfaktor	S. 70
Tabelle 14:	Analyse-Ergebnisse für die CT- Studien in Bezug auf den Impaktfaktor	S. 70
Tabelle 15:	Analyse-Ergebnisse der MRT-Studien in Bezug auf den Impaktfaktor	S. 71
Tabelle 16:	Impaktfaktoren und Evidenzlevel der Studien	S. 90
Tabelle 17:	Ausschlussgründe für gefundene Studien	S. 93
Tabelle 18:	Levels of Evidence	S. 98

### **13. VERZEICHNIS DER AKADEMISCHEN LEHRER**

Adamkiewicz, Bartsch, Baum, Baumann, Becker, Berger, Bette, Bösner, Brehm, Cetin, Czubayko, Daut, Del Rey, Donner-Banzhoff, Fendrich, Feuser, Görg, Gress, Hadji, Hertl, Höffken, Hofmann, Holst, Hoyer, Hundt, Jacobsohn, Jerrentrup, Kann, Kill, Kinscherf, Kircher, Klose, Koolman, Köhler, König, Lill, Lohoff, Mahnken, Maier, Maisch, Moll, Moosdorf, Mueller, Mutters, Neubauer, Nimsky, Oertel, Olbert, Pagenstecher, Peterlein, Rausch, Renz, Richter, Rosenow, Rothmund, Röhm, Ruchholtz, Sahmland, Schäfer, Schieffer, Schu, Schütz, Seifert, Seitz, Sekundo, Sesterhenn, Sevinc, Sitter, Sommer, Steiniger, Vogelmeier, Vogt, Wagner, Weihe, Werner, Westermann, Wilhelm, Wulf, Wrocklage, Ziring

## 14. DANKSAGUNG

Mein besonderer Dank gilt meinem Doktorvater, Herrn Priv.-Doz. Dr. Helmut Sitter, für die Überlassung des Themas und seine Unterstützung und Betreuung. Ohne seinen akademischen Rat in inhaltlichen sowie in methodischen Fragen wäre diese Arbeit nicht möglich gewesen. Auch für die Hilfestellung bei der Formatierung bin ich ihm sehr dankbar. Seine Anregungen und kritischen Kommentare haben maßgeblich zum Gelingen dieser Arbeit beigetragen.

Außerdem danke ich Herrn Prof. Dr. med. Hauke Lang (Universität Mainz) ganz herzlich für die Durchsicht der Arbeit und seine konstruktive Kritik. Insbesondere für die klinischen Bezüge hat mir seine Expertise ungemein geholfen.

Bedanken möchte ich mich ebenso bei Frau Barbara Gstrein für ihre Hilfestellung bei der Literatursuche und /-beschaffung.

Ich danke außerdem meinen Freunden zu Hause und in Marburg dafür, dass sie stets ein offenes Ohr für mich haben, mich immer unterstützen und in allen Lebenslagen für mich da sind. Ich danke euch für eure Unterstützung und euren Rückhalt in dieser wichtigen Phase meines Lebens!

Meinem Freund Michael danke ich für seine großartige Hilfe in allen größeren und kleineren Notsituationen. Ob bei Formatierungsfragen, Problemen bei der Formulierung oder Statistikrätseln - ich konnte mich jederzeit auf ihn verlassen und habe durch ihn immer wieder neue Motivationsschübe erhalten. Vielen Dank für deine Geduld und deinen unerschöpflichen Optimismus!

Auch meinem Bruder Robert möchte ich dafür danken, dass er mich immer aufmuntert, wenn ich es brauche und immer Zeit für meine Sorgen hatte, in einer Zeit in der er selber nicht sorgenfrei war. Du bist großartig!

Nicht zuletzt möchte ich mich ganz besonders bei meinen Eltern Karin und Emil Pado bedanken, die mir mein ganzes Leben lang mit Rat und Tat zur Seite standen. Ich danke euch dafür, dass ihr mir meine Auslandsaufenthalte und das Medizinstudium ermöglicht habt, sowie für eure liebevolle Unterstützung und euer Vertrauen in mich. Ohne euch wäre ich heute nicht die, die ich bin. Ihr seid die besten Eltern, die man sich wünschen kann!